UAB
Universitat Autònoma de Barcelona

# Essays on GDP per Capita Prediction, Conflict-Poverty Dynamics, and Foreign Aid for Natural Resources

**Chanon Techasunthornwat**

# Essays on GDP per Capita Prediction, Conflict-Poverty Dynamics, and Foreign Aid for Natural Resources

*A Doctoral Thesis Submitted in Partial Fulfillment of the Requirements for the Degree of*

Doctor of Philosophy in Economics

**by**

Chanon Techasunthornwat

**Supervisor:**

Hannes Mueller

International Doctorate in Economic Analysis

Universitat Autònoma de Barcelona and Barcelona School of Economics

September, 2024

# Acknowledgments

I would like to express my deepest gratitude to my supervisor, Hannes Mueller, whose patience, kindness, and honesty have been invaluable throughout my academic journey. His guidance has shaped not only my research but also my approach to teaching and supervising future researchers.

My sincere thanks extend to the members of my thesis tribunal, who generously sacrificed their valuable time to read and provide insightful suggestions for this work. Their contributions have significantly enriched this thesis.

The completion of the second chapter was made possible by the support, insightful comments, and suggestions from Samuel Freije-Rodriguez, Michael Woolcock, Alexandru Cojocaru, Minh Cong Nguyen, Patrick Barron, and the participants of the Poverty, Fragility, and Conflict Webinar, for which my coauthor and I are grateful.

I would also like to express my sincere gratitude to Thammasat University for the financial support that made this study possible.

To my family, especially my parents, I offer heartfelt thanks for their support and understanding. Their encouragement has been a cornerstone of my journey.

I would like to acknowledge Anan Pawasutipaisit, Wanwiphang Manachotphong, Nattapong Puttanapong, Thorn Pitidol, and Somchai Jitsuchon for playing a pivotal role in nurturing my intellectual development during my master's program at Thammasat University and the early stages of my career at the Thailand Development Research Institute. Their support and encouragement were instrumental in helping me embark on my Ph.D. journey.

Special thanks to my colleagues and the staff of the IDEA program for making this journey a smooth and enriching experience. In particular, Àngels and Mercè have been exceptional in their support, guiding us from Day 0 to the very end with flawless assistance.

Finally, to Zhang Boyao, I am forever grateful for your unwavering support, companionship, and the balance you provided between thoughtful discussions and much-needed recreation.

All errors and shortcomings in this thesis remain my own.

# Abstract

This doctoral thesis consists of three chapters in development economics, centered around improving the understanding of relationships among socioeconomic factors in developing countries, where data availability and working conditions for policymakers are often challenging. The first chapter addresses a crucial question in empirical development economics: how to assess the living standards of populations in developing countries where official data are scarce. The following two chapters present theoretical frameworks to explore the relationships between conflict and poverty, as well as foreign aid and natural resources, supported by empirical tests of the proposed theories. Through these new insights, this thesis aims to modestly contribute to the broader understanding of these complex issues.

The first chapter addresses subnational GDP data gaps, particularly in Southeast Asia, by leveraging satellite imagery and machine learning techniques. It introduces a transfer-learning toolbox designed for estimating subnational GDP per capita, particularly suited to data-constrained environments due to its strong performance, feasibility, and low data access costs—crucial factors in the context of developing countries. The model incorporates a comprehensive set of daytime and nighttime satellite imagery features and demonstrates competitive performance, achieving an out-of-sample R-squared of 0.77 in predicting GDP per capita. Lasso regression further confirms the external validity of the method, revealing minimal dependence on area or population size. Additionally, a parsimonious model, relying solely on daytime features and capable of producing predictions within three weeks after the target period, delivers an acceptable out-of-sample R-squared of 0.70. When compared to a more data-intensive method employed in a previous study, this approach provides highly competitive GDP predictions, with an R-squared of 0.94. However, for the more complex task of predicting GDP growth, the proposed method exhibits underperformance, indicating potential for improvement through the incorporation of advanced deep learning techniques and a larger training dataset.

The second coauthored chapter explores the relationship between poverty and conflict using both macro-level and regional data, with a detailed case study of Uganda. This analysis builds on a substantial and growing body of literature that documents the

severe impacts of conflict on health and economic outcomes. Drawing on this evidence, we develop a statistical framework to trace the cumulative long-term effects of armed conflict on poverty, which we term “conflict debt.” Our findings reveal that contemporary conflicts generate a conflict debt, which is only gradually alleviated over time. The empirical model not only successfully captures cross-country aggregate poverty trends but also reflects regional cross-sectional variations. A key feature of the model is its recognition that armed conflict can both impede poverty reduction and, once resolved, trigger significant catch-up effects, as demonstrated by the data. However, in countries most severely affected by conflict, recurrent cycles of violence significantly obstruct poverty recovery. Even our most conservative estimates suggest that, in the absence of conflict debt, these countries and regions would experience a 5-10 percentage point reduction in poverty rates.

The final chapter investigates the hypothesis that wealthy donor countries without significant natural resources, such as oil and gas, may provide financial aid to resource-rich but economically poorer countries to secure access to these resources. This chapter introduces a theoretical framework to examine the “aid-for-resource” motivation alongside altruistic motives. The theory outlines three key mechanisms influencing aid: the valuation of the recipient’s natural resources, the recipient’s income, and aid from other donors. Empirical tests, utilizing exogenous variation in global fossil fuel prices, confirm the direct impact of resource values on aid commitments. Specifically, a 1% increase in resource value leads to a 1.1 percentage point greater increase in development aid from high-resource-demand donors compared to lower-demand donors. Donor-specific aid elasticities with respect to resource values are calculated, revealing that countries with high demand for resource imports—such as the United States, the United Kingdom, France, Germany, Japan, and South Korea—demonstrate the highest aid elasticities to natural resource values, consistent with the behavior of resource-motivated donors as predicted by the theory. Interestingly, Sweden, despite its relatively modest demand for resources, also displays high elasticity.

# Contents

# Chapter 1

# Predicting Subnational GDP per Capita in Developing Countries Using Machine Learning and Satellite Images

## 1.1 Introduction

GDP data at the subnational level is often scarce, particularly in developing countries. As of 2020, many African and Middle-Eastern countries do not possess a subnational GDP database. Moreover, certain regions in Latin America and Southeast Asia, such as Paraguay and select provinces in Argentina, along with Laos, Myanmar, Cambodia, and Timor-Leste, also lack such data (Wenz et al., 2023). The absence of this data can have significant implications in developing countries, posing challenges for policymakers in prioritizing development efforts and hindering endeavors to attract investment and promote economic growth.

To address the data gap and enhance data reliability, remote-sensing data is a valuable resource. These publicly available databases have improved thanks to advancements in computer science, engineering, and geography, making them beneficial for economic research in the era of data revolution (Donaldson and Storeygard, 2016).

One increasingly crucial area of economic research reliant on remote-sensing data is the estimation of economic outcomes in regions with limited data availability. On the demand side, the Statistical Commission of the United Nations—an international organization dedicated to promoting global development goals—established the UN Committee of Experts on Big Data and Data Science for Official Statistics (UN-CEBD). This committee aims to explore the potential of big data for tracking and reporting on sustainable development goals. It also provides guidance to national statistical offices on integrating big data into their national statistical methods.

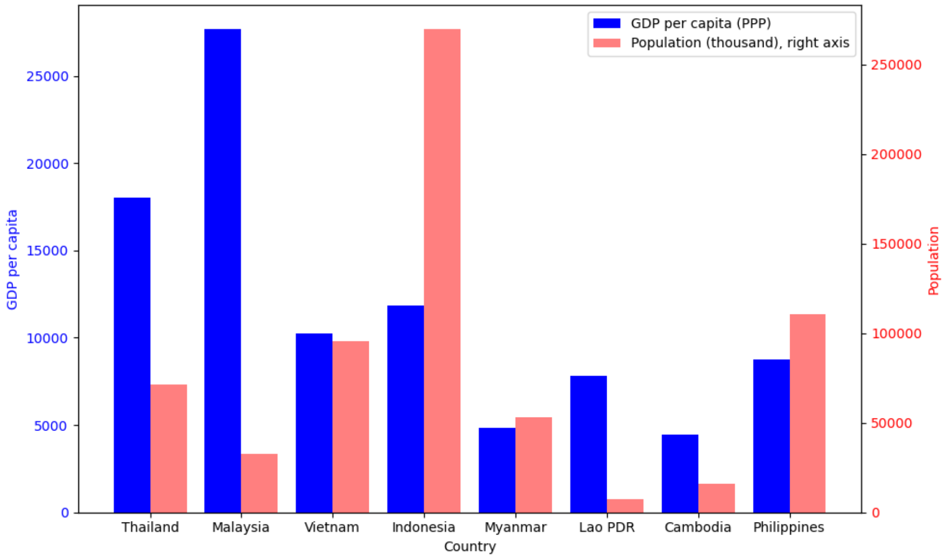
On the academic side, researchers from various international organizations are collaborating on remote-sensing data projects for development, including UNICEF (Tingzon et al., 2019), ADB (ADB, 2020), and the World Bank and IMF (Babenko et al., 2017; Hernandez et al., 2017; Beyer et al., 2022). To contribute to the policy community, this study introduces a toolbox for estimating subnational GDP per capita using satellite images and machine learning.

The area of interest in this study is South East Asia, which is known for its diverse living standards. This region includes countries with varying economic conditions, such as those in the midst of civil conflicts or experiencing economic slowdown at the middle-income stage, resource-rich and fast-growing nations, as well as middle- or high-income countries with stable growth. Figure 1.1 illustrates the diversity of the region in terms of GDP per capita and population size. Despite this diversity, the shared proximity and geographical characteristics make it an excellent choice for developing a regional model.

Another crucial reason for selecting South East Asian countries for this experiment is the limited availability of data in the region. Our survey revealed that Vietnam lacks a centralized database for provincial GDP, with most figures coming from local authorities' annual reports. As of September 2023, Thailand's most recent data is from 2021, and Laos, Cambodia, and Myanmar have no available data. Working in this data-poor environment not only highlights the data gap issue but also directly benefits policymakers. We have also included states and counties from India and China to increase the sample size.[1]

[1]The high number of Chinese counties may dominate the sample. However, we have implemented splitting rules for training and validation to balance the sample across countries, mitigating this issue.

Figure 1.1: GDP per capita and population 2019




Source: World Development Indicators database, the World Bank

A distinctive aspect of this study is its focus on the feasibility of implementation due to data constraints, making the propose framework accessible to national statistical offices in developing countries. This is achieved by using a small number of databases and a moderate computational demand for the estimation. Consequently, the approach revolves around using only two main sources of satellite images and transfer-learning method, making it suitable for small-scale data environments.

This chapter proposes a regional framework for predicting GDP per capita using satellite images, incorporating both daytime and nighttime data, through the transfer-learning method. A notable toolbox, known as *"Nowcasting"*, utilizes all up-to-date data and images to predict the current outcome variable. The results are promising, with an out-of-sample R-squared value of 0.77 for predicting GDP per capita in 2019. To provide a benchmark for this method, we follow the same steps to predict GDP values, a common label in the literature, and achieve a more competitive accuracy with an out-of-sample R-squared of 0.94.

For implementation, the speed of prediction depends on the availability of data sources. Daytime images from Sentinel-2 are updated weekly, typically three weeks after the target period, while nighttime VIIRS data is published annually, around six months after the year-end. Policymakers may face a trade-off between applying the complete model, which offers high accuracy but delayed results, and using the quick prediction method, which relies only on daytime features and provides reasonably accurate estimated GDP per capita with an out-of-sample R-squared of 0.70 when a timely forecast is required.

One limitation of the proposed toolbox is its accuracy in predicting growth rates in GDP per capita and GDP, particularly in scenarios involving extreme economic changes and human migration. Compared to benchmark results, the toolbox demonstrates lower precision. However, as this study is among the few to examine both GDP per capita and growth rates, this underperformance may be mitigated in future work by employing more advanced neural network training methods suited for time-series data, and by incorporating a larger dataset to enhance the training of the Random Forest model.

This article comprises six sections. It begins by discussing related studies in the next section. Following that, a package of datasets used for the study is introduced with descriptive statistics. Section 1.4 provides an explanation of the method of transfer learning and outlines the steps to reproduce the results shown in Section 1.5. Within Section 1.5, a detailed discussion is presented on how well the proposed toolbox can perform under different specifications, a comparison to the benchmark, and its external validity (Section 1.5.1). Additionally, this section explores the important daytime features in the model for predicting the outcome (Section 1.5.2), the results with varying training sets (Section 1.5.3), the robustness of the results by country (Section 1.5.4), and the performance of predicting changes in GDP per capita (Section 1.5.5). The final Section 1.6 highlights important results and discusses the implementation of the method for policymakers.

## 1.2 Related Literature

The related literature of this study involves research estimating economic measures using unconventional data. The first group of literature is related to forecasting macroeconomic factors using big data. The wave of big data use to predict current economic situations was initiated by Choi and Varian (2012). They introduced Google Trend time series generated by search queries from the website to predict sales of auto parts, claims for unemployment benefits, and visitors to Hong Kong. Examples of subsequent studies include Bok et al. (2018) and Buell et al. (2021). However, the performance in predicting GDP may vary across income groups. For example, Narita and Yin (2018) found that the prediction is less accurate among middle-income emerging countries. This result raises concerns about the strength of the link between internet data and the real economy, as well as the external validity of the method itself.

The second group of the literature is more closely aligned with this study, employing remote-sensing data as predictors. Nighttime light data is one of the most common datasets in the economic literature, initially introduced by Sutton et al. (2007) and Henderson et al. (2012). They demonstrate that luminous values have a strong correlation with economic development. Many economic studies have followed suit (Chen and Nordhaus, 2011; Hodler and Raschky, 2014; Dreher and Lohmann, 2015; Henderson et al., 2017; Wang et al., 2019; Gibson et al., 2020; Beyer et al., 2022). One caveat of nighttime light data is that its connection with economic development is minimal in less developed areas where brightness is low (Jean et al., 2016).

Newer features for predicting economic indicators are derived from daytime satellite images. When combined with machine learning techniques, these images offer a wealth of spatial characteristics, even in less developed regions. Many researchers have proposed frameworks that integrate this data with various machine learning methods, resulting in highly accurate predictions. It's worth noting that the outcome variables estimated in these studies encompass wealth, GDP values, or population (Nordhaus and Chen, 2016; Kummu et al., 2018; Jean et al., 2016; Yeh et al., 2020; Khachiyan et al., 2022), while another group focuses on estimating poverty rates or human development indicators (Babenko et al., 2017; Kummu et al., 2018; Tingzon et al., 2019; Jean et al., 2016; ADB, 2020; Puttanapong et al., 2022).

The outcome measures in the former group reflect the overall economic activities within the areas captured by the images, closely aligning with the input images when the research question intends to estimate the value of assets, economic activity, or population within this area. However, these measures may somewhat represent less of human living standards. On the other hand, the predicted outcomes in the latter group, such as poverty rates and development indicators as in Kummu et al. (2018), are more closely associated with reflecting the living standards of the human population in the captured areas. Nevertheless, it's important to acknowledge that definitions of people living in poverty are not universally applied, and their relevance can vary across countries. For instance, some middle-income countries may not exhibit high or varying poverty rates when measured against the global poverty line, unlike lower-income countries. In the context of this study, the choice is made to predict GDP per capita, a more widely used measure of living standards and economic development that allows for comparisons across regions.

A few previous studies are discussed in detail due to their key concepts being related to this chapter. The first paper, Khachiyan et al. (2022), published in *AER Insights*, serves as a benchmark for this study. The authors introduce a method for predicting income at the block-group level by directly training a neural network using daytime satellite images. The dataset used in this research consists of satellite data extracted from 7-band Landsat images, specifically capturing urbanized areas during the months of May through August. Income and population data were collected from the US Census and Census Block data, along with data from the American Community Survey (ACS). These data were then interpolated from block-level to image-level for analysis. The method is supported by enriched datasets containing high-resolution label variables, block-level income, and detailed 7-band imagery. This setup indicates that the computational demands are significant, making it nearly impossible to replicate the training in data-limited environments, such as those found in many developing countries. Nevertheless, we regard it as a high-performing benchmark for our study.

In terms of model performance, the study achieved impressive results on 5.76 sq.km. imagery, with a validation $R^2$ of 0.84, serving as the benchmark for our method.[2] Additionally, the study conducts change prediction, yielding an $R^2$ of 0.396—a noteworthy accomplishment considering the absence of a prior benchmark in the existing literature.[3]

To assess the robustness of the findings, various tests were executed. When utilizing the less-detailed 3-band RGB imagery, the $R^2$ reduced modestly by 0.04, indicating a slight decrease in predictive performance. Furthermore, experiments with higher-resolution images did not yield any meaningful improvements in model fit across different specifications. These findings provide valuable insights into the efficacy of the proposed model and its adaptability to different data configurations.

A seminal paper by Jean et al. (2016) published in *Science* introduces a transfer learning approach tailored for developing-country contexts. Their study harnessed household surveys from the Living Standards Measurement Study (LSMS) and satellite data spanning regions in Nigeria, Tanzania, Uganda, Malawi, and Rwanda.

The data collection involves nighttime images from the Defense Meteorological Satellite Program's Operational Linescan System (DMSP-OLS) and daytime images obtained from Google Static Maps. The transfer learning method comprises several essential steps. It begins with a pretrained convolutional neural network (CNN) on the ImageNet dataset, which is then fine-tuned to predict nighttime light intensities based on the input daytime satellite imagery, resulting in a retrained CNN. Principal Component Analysis (PCA) was employed to reduce the feature dimensions to 100, streamlining the data for analysis. Finally, ridge regression was applied, using the features extracted from the daytime images by the retrained CNN, to predict average economic outcomes derived from the surveys.

The results of their study are impressive, with an R-squared value of 0.56 for predicting household consumption per capita in Tanzania and an even higher R-squared value of 0.75 for predicting asset values in Rwanda. These findings highlight the effec-

[2] This represents the model's performance without additional features collected from the census. The model, incorporating these census features, performs better with a validation $R^2$ of 0.90. However, this highly demanding model is incomparable to our method.

[3] Yeh et al. (2020) also predict changes in wealth, achieving performance scores ranging from 0.35 to 0.51 in terms of R-squared.

tiveness of their transfer learning approach in data-limited contexts, such as those found in many developing countries.

Rolf et al. (2021) share the same spirit as this chapter. They also employ satellite imagery to generate a universal set of features using unsupervised machine learning. These versatile features can be applied in linear regression models to analyze various outcomes, ranging from predicting forest cover and elevation to population density and housing prices. Even though their performance in predicting household income, an essential economic variable, is not particularly strong, the focus on low computational costs and the potential for implementation by statistical offices in developing nations deserves further exploration and emphasis in the literature.

Building on the existing literature, this study contributes by demonstrating the application of transfer learning for the economically significant and specific task of predicting GDP per capita in developing countries. The data-intensive method proposed by Khachiyan et al. (2022), which is nearly impossible to implement in data-constrained environments, will be utilized as a gold standard benchmark for performance comparison. The newly developed toolbox aims to balance adequate performance with practical feasibility and implementation costs, particularly in terms of data access, thereby supporting developing countries with limited data and computational resources for machine learning initiatives.

## 1.3 Data

This section presents data sources that underpin the methodology. The first source purposes for defining the sampling frame of our satellite imagery dataset. Subsequently, we discuss the label data, specifically GDP per capita, showing how it's prepared for the machine learning exercise. Lastly, we describe two remote-sensing datasets, covering both daytime and nighttime images, as the sources of predictors in the machine-learning process.

It's important to note that our sample for this analysis is restricted to seven countries: five countries from South East Asia, India, and China. Within South East Asia, the included countries are Indonesia, Malaysia, Vietnam, the Philippines, and Thailand, all of which provide subnational GDP per capita data. Additionally, India and China are incorporated into the sample due to their geographic proximity to the South East Asia, enhancing the size of observations available for cross validation in machine learning process.

In our implementation strategy, we prioritize data choices that minimize the financial burden on national statistical offices, especially in developing countries where expertise in remote sensing data, coding, and machine learning might pose implementation challenges. To streamline the process, we opt for a simplified set of features for prediction, encompassing both nighttime and daytime satellite images. This approach aims to leverage the complementary nature of these data sources while ensuring free access and cost-effective implementation.

### 1.3.1 The GHS Settlement Model grid

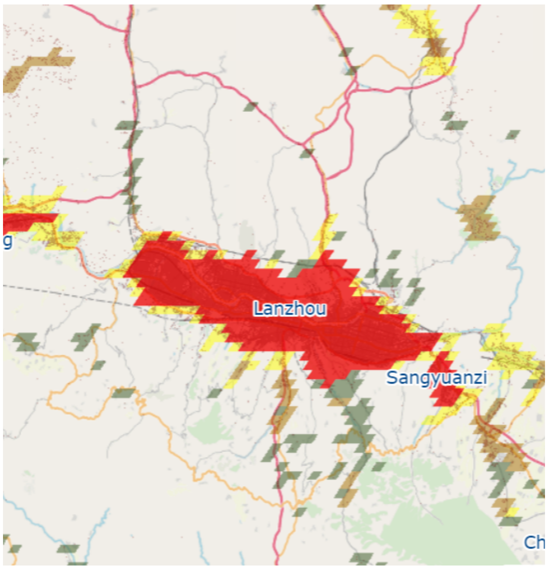
The GHS Settlement Model grid (GHS-SMOD) is one of raster datasets from the Global Human Settlement Layer (GHSL) project. The GHS-SMOD dataset delineates the spatial boundaries of human settlements, thereby offering the distribution of urban and rural areas at a 1-km resolution. Furthermore, the dataset provides the density of human settlements, a key metric that offers perspectives on the degree of urbanization. Specifically, this study relies on the *GHS-SMOD_GLOBE_R2023A* version of the dataset, which contains the raster of global urban areas in 2020 (Commission and Centre, 2023; Schiavina et al., 2023).

The GHS Settlement Model grid (GHS-SMOD) is constructed from two precursor models: the GHS population spatial raster dataset (GHS-POP) and the GHS built-up surface spatial raster dataset (GHS-BUILT-S). GHS-POP is instrumental in estimating the residential population within each grid cell, employing a temporal span from 1975 to 2030. This estimation process relies on raw global census data that has been harmonized by the Center for International Earth Science Information Network (CIESIN) for the Gridded Population of the World, version 4.11 (GPWv4.11). The distribution of the estimated population across grid cells is executed using GHS-BUILT-V, a modified iteration of GHS-BUILT-S. On the other hand, GHS-BUILT-S defines the distribution of built-up (BU) surfaces within the same temporal intervals. The underlying data for this model is garnered from satellite imagery collected by the Landsat and Sentinel platforms. This mixture of data sources and processing steps forms the GHS-SMOD, enabling a detailed portrayal of human settlement patterns and urbanization trends.

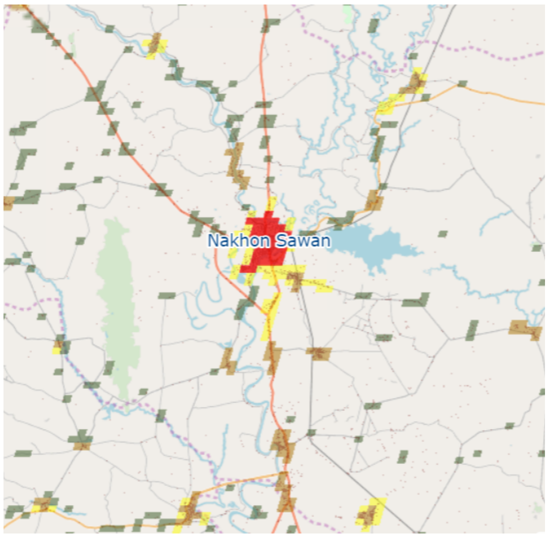
The GHS-SMOD dataset categorizes global cells into three distinct classes: Rural grid cells (Class 1), Urban Cluster grid cells (Class 2), and Urban Centre grid cells (Class 3). As shown in Figure 1.2, two illustrative examples of Lanzhou in China and Nakhon Sawan in Thailand highlight the spatial distribution of these cell classes on the map. In this study, Class 1 cells are selectively excluded from the sample, while the remaining classes are retained for the purpose of estimating GDP per capita. The distribution of these cells in each country is presented in Table 1.1. The distribution of rural cells varies significantly across countries, ranging from 79.6 percent in India to 96.4 percent in Malaysia. Consequently, the proportion of cells included in the study falls within the

Figure 1.2: Examples of GHS-SMOD classes on the map

a) City of Lanzhou, China




b) Nakhon Sawan, Thailand




Notes: Green cells represent Rural grid cells *(Class 1)*, while yellow and brown cells correspond to Urban Cluster grid cells *(Class 2)*, and red areas denote Urban Centre grid cells *(Class 3)*.

range of 3.6 to 20.4 percent. This variability in rural-urban distribution among countries highlights the diverse subregions under examination.

The application of this sampling frame to image datasets is based on a fundamental assumption, namely, that the primary determinant of variations in GDP per capita over time or differences across regions lies in the extent of structural transformation. Economic development within a region is typically realized when its labor force transitions from traditional or low-productivity sectors to formal or high-productivity sectors predominantly situated in urban areas. Hence, more economically developed regions can be detected by their larger and more advanced urban areas. This underlying premise underscores the pivotal role played by structural shifts in fostering economic growth and development across regions (Herrendorf et al., 2014).

A comparable approach to our methodology is the importance sampling method, which selectively samples only the relevant observations to explain the outcome. In contrast, employing uniform sampling in this context would lead to numerous empty observations, such as images of deserts or remote mountainous regions where economic activity is minimal (Jean et al., 2016; Price and Atkinson, 2022; Khachiyan et al., 2022).

Table 1.1: Share of cells in GHS-SMOD

| Class | IDN | IND | MYS | PHL | THA | VNM | CHN |
|---|---|---|---|---|---|---|---|
| Class 1 | 93.9 | 79.6 | 96.4 | 82.8 | 94.9 | 83.4 | 93.5 |
| Class 2 | 4.8 | 18.0 | 2.4 | 14.9 | 4.1 | 14.4 | 5.4 |
| Class 3 | 1.3 | 2.4 | 1.2 | 2.3 | 1.0 | 2.2 | 1.1 |

Source: Author's tabulation from GHS-SMOD.
Notes: *Class 1* corresponds to rural grid cells, *Class 2* represents urban cluster grid cells, and *Class 3* signifies urban center grid cells.

### 1.3.2 Subnational GDP per capita

As previously mentioned, our study involved the collection of subnational GDP per capita data from five countries within the ASEAN region, China, and India, sourced from their respective national statistical offices. To ensure the highest granularity of data possible, we selected the finest subnational level available. Consequently, the datasets for GDP per capita encompass six countries, all measured at the ADM-1 level, which corresponds to the province or state level. Notably, in the case of China, we were able to access data at an even more detailed level, specifically the county level (ADM-3). In total, our dataset comprises 2,803 distinct labels per year. The labels from 2018 and 2020 are gathered to perform a nowcasting exercise.

Our data processing methodology involves converting the GDP per capita values from local currency into real values as of the year 2019. To ensure cross-national comparability, we further adjust these values using purchasing power parity (PPP) conversion factors. By applying these conversions and adjustments, the data accurately reflect the relative economic well-being of individuals across nations and facilitate a more robust analysis of the factors influencing living standards on an international scale.

Table 1.2 provides a summary of descriptive statistics for adjusted GDP per capita 2019 in logarithmic form. Notably, Malaysia exhibits the highest average GDP per capita within the sample, whereas the Philippines and Vietnam display the lowest average values.[4] Furthermore, Thailand and China stand out as the two countries with the most pronounced income inequality across their regions, as indicated by the high standard deviation among the countries considered. Interestingly, China demonstrates significant variability in subnational income, encompassing both the wealthiest and the poorest subregions. This diversity contributes to China's position as one of the countries with the highest variance in the sample, attributable to its substantial number of subregions.

The choice of using GDP per capita as the target variable for prediction, rather

[4]It is important to note that these statistics are not population-weighted, thus representing an approximation rather than exact national GDP per capita figures.

than absolute GDP values as employed in some prior studies, is rooted in its significance as a proxy for measuring living standards. However, as part of our research methodology, we conducted an additional experiment where we trained the model using GDP values. This decision was made to facilitate a meaningful benchmarking exercise, enabling us to compare our approach with those of other studies.

Table 1.2: Descriptive statistics of GDP per capita (in log) 2019

| ISO3c | Count | Mean | SD | Min | 25pct | 50pct | 75pct | Max |
|---|---|---|---|---|---|---|---|---|
| CHN | 2,564 | 9.29 | 0.72 | 7.22 | 8.77 | 9.17 | 9.74 | 12.11 |
| IDN | 34 | 9.32 | 0.55 | 8.33 | 8.97 | 9.30 | 9.42 | 10.94 |
| IND | 33 | 8.90 | 0.54 | 7.65 | 8.51 | 9.06 | 9.24 | 9.94 |
| MYS | 15 | 10.12 | 0.54 | 9.10 | 9.77 | 10.06 | 10.43 | 11.30 |
| PHL | 17 | 8.85 | 0.45 | 7.95 | 8.64 | 8.73 | 9.13 | 10.08 |
| THA | 77 | 9.25 | 0.80 | 8.27 | 8.64 | 9.07 | 9.62 | 11.60 |
| VNM | 63 | 8.82 | 0.41 | 8.08 | 8.55 | 8.73 | 8.99 | 9.88 |

Sources: National statistical offices of the respective countries, adjusted by the author to represent values in international dollars for the year 2019, taking into account country-level GDP deflators and the PPP conversion factor.

### 1.3.3 Nighttime images

The initial set of features employed in our machine learning framework comprises nighttime images obtained from the Visible and Infrared Imaging Suite (VIIRS) Day Night Band (DNB). This particular dataset offers a significant enhancement in low-light imaging capabilities when contrasted with the Defense Meteorological Satellite Program featuring the Operational Linescan Sensor (DMSP-OLS), which has traditionally been popular among economists (Elvidge et al., 2017). Gibson et al. (2020) highlight a series of advantages offered by the VIIRS dataset in comparison to the Defense Meteorological Satellite Program with the Operational Linescan Sensor (DMSP-OLS). One key distinction lies in the intended purpose of these systems, as DMSP-OLS was primarily designed to support the daily weather forecasting needs of the United States Air Force, rather than for research purposes, unlike VIIRS. Moreover, DMSP-OLS lacks onboard calibration mechanisms to ensure the consistency of readings over time, a feature that VIIRS incorporates. Additionally, VIIRS boasts a notably superior dynamic range, enabling it to capture scenes ranging from the brightest daytime landscapes to extremely dim nighttime environments illuminated by a quarter moon. With a minimum and maximum value range spanning from -1.5 to 340,573, VIIRS effectively prevents image saturation in the presence of intense brightness.

In this chapter, we make use of the VIIRS time series dataset in the form of the annual composite version 2, known as Annual VNL V2. This dataset is constructed by aggregating monthly observations of cloud-free average radiance. It spans from the year 2012 to the most recent available data. The satellite captures nighttime imagery around 1:30 am, minimizing geolocation errors and ensuring data accuracy. The processed nighttime images boast an approximate resolution of 465*465 meters at the equator, offering a high level of detail for our analysis. However, it's worth noting that this dataset is published annually, typically released about six months after the end of the year. This publication frequency could be considered a limitation, particularly when our aim is to develop a nowcasting toolbox for timely predictions of GDP per capita, as it may introduce a lag in data availability for real-time forecasting.

Within each subregion, we compile two groups of basic statistical metrics derived from the nighttime images: firstly, the average DNB value of pixels located within the urban areas, as defined by GHS-SMOD (*urban_ntl*), its standard deviation (*urban_std*), and size (*urban_size*); secondly, the average DNB value encompassing the entire subregion (*total_ntl*), its standard deviation *total_std*, and size *total_size*. Three features of nighttime images are summarized in Table 1.3, utilizing images from the year 2019.[5] Notably, Malaysia stands out as the brightest urban area and country, a consistency that aligns with GDP per capita trends. It's worth noting that the sizes of the sampled subregions exhibit variation, with China featuring the smallest average size, as expected given its county-level data. In fact, the smallest subregion is also located in China, while the largest subregion is an Indian state. This diversity in subregion sizes contributes to the method's robustness and ability to generalize across regions with varying spatial extents.

Table 1.3: Descriptive statistics of nighttime light values 2019

| ISO3c | urban_ntl | | | | total_ntl | | | | total_size | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | mean | std | min | max | mean | std | min | max | mean | std | min | max |
| CHN | 1.31 | 0.82 | 0.00 | 4.34 | 0.73 | 0.87 | 0.00 | 4.63 | 9.05 | 1.26 | 3.95 | 13.99 |
| IDN | 1.27 | 0.56 | 0.37 | 3.54 | 0.40 | 0.66 | 0.01 | 3.53 | 11.95 | 1.20 | 8.02 | 14.21 |
| IND | 1.28 | 0.62 | 0.52 | 3.42 | 0.68 | 0.73 | 0.03 | 3.29 | 12.21 | 1.88 | 6.36 | 14.39 |
| MYS | 2.86 | 0.54 | 2.10 | 4.12 | 1.48 | 1.11 | 0.15 | 4.13 | 10.24 | 2.06 | 6.10 | 13.27 |
| PHL | 0.77 | 0.71 | 0.20 | 3.27 | 0.38 | 0.77 | 0.04 | 3.29 | 11.16 | 0.85 | 7.97 | 11.77 |
| THA | 1.88 | 0.42 | 1.21 | 3.27 | 0.71 | 0.67 | 0.04 | 3.16 | 10.09 | 0.87 | 7.60 | 11.63 |
| VNM | 1.26 | 0.81 | 0.20 | 4.48 | 0.69 | 0.73 | 0.02 | 3.32 | 9.89 | 0.76 | 8.32 | 11.31 |

Source: Author's tabulation from VIIRS by Elvidge et al. (2017).
Notes: *urban_ntl* denotes the average DNB value within the urban area, while *total_ntl* represents the average value for the entire subregion. *total_size* corresponds to the subregion's area size. Note that all features undergo a standardization for the analysis. Descriptive statistics of more nighttime features is in Table 1.10.

[5] The rest of the nighttime features are in Table 1.10.

### 1.3.4 Daytime images

We utilize daytime images from Harmonized Sentinel-2 MSI Level-2A data, which boasts a spatial resolution of 10 meters for the visible Red-Green-Blue (RGB) bands, delivering comprehensive Earth surface details. The satellite's revisit frequency is approximately 10 days at the equator, enabling temporal analysis and change detection. The sensor's spectral bands provide precise insights into land cover, vegetation health, and water quality, rendering it an indispensable tool for tasks such as crop monitoring, forest management, and environmental studies. Furthermore, Level-2A processing encompasses radiometric and geometric corrections, ensuring calibrated images with accurate georeferencing that represent true surface colors. This process enhances their suitability for scientific research and operational applications, compared to the Level-1C process.

The speed of data publication is a crucial feature of the Sentinel-2 system. With new images being uploaded at a high frequency, it enables near real-time access to data. This means that the nowcasting of GDP per capita for a given year can commence immediately after the year ends, allowing for timely and up-to-date economic analysis and decision-making based on the latest available information.

However, obtaining the corrected surface reflection from this modified dataset comes at a cost—the earliest available images in the sample countries date back to December 1st, 2018.[6] Consequently, to create daytime images that accurately represent economic development in 2018, we must extract data from the period spanning December 1st, 2018, to February 28th, 2019. On the contrary, for images capturing the state of affairs in 2019 and later years, the sampled time frame is extended to span from September 1st to December 31st of the corresponding year.

Our choice of selecting the period around the end of each year is driven by how GDP is measured. Given that GDP encompasses all economic activities within the calendar year, imagery captured at the year's end aligns more closely with this measurement. This is because changes in man-made infrastructure are akin to stock variables, reflecting the cumulative effects of economic activity over time. Therefore, the optimal moment for sampling images that can accurately depict economic activity in 2019 is on December 31st, 2019. In practice, we find it necessary to extend the observation period backward to September, a strategic move that not only increases the sample size but also allows us to generate an average composite of each image, providing a more comprehensive representation of economic development within the year.

In each year of our study, we collected images from urban areas within the sampled

[6]These images are sourced from Google Earth Engine, which provides free access and a cloud-free image generation process.

subregions as outlined by the GHS-SMOD dataset. For subregions where urban areas contained more than 500 images, we implemented a random selection process to acquire 500 images per urban area. Each image had a fixed size of 224x224 pixels, providing coverage for approximately 5.02 square kilometers of the Earth's surface. In total, our dataset consisted of 220,264 images, with an average of 78.6 urban images per subregion. Notably, a substantial majority, 77 percent, of these images originated from Chinese counties, while the Philippines exhibited the highest average number of images per subregion, with an impressive 419 images on average. For a comprehensive breakdown of these statistics, please refer to Table 1.4, which provides details on the number of images downloaded from the Sentinel-2 dataset.

Lastly, we process the collected daytime images by excluding those with more than 50 percent transparent pixels, typically caused by cloud masking. It's worth noting that this issue is more prevalent among the 2018 images (nearly 11 percent) due to the shorter sampling window of four months, but it is less common in 2019 (0.8 percent). As a result, 2.2 percent of the entire image collection is omitted.

Table 1.4: Daytime images downloaded from Sentinel-2 per year, by country

| ISO3c | Image No. | Share | Images per subregion |
|---|---:|---:|---:|
| CHN | 170,459 | 77.4 | 66.5 |
| IDN | 10,570 | 4.8 | 310.9 |
| IND | 12,689 | 5.8 | 384.5 |
| MYS | 2,393 | 1.1 | 159.5 |
| PHL | 7,120 | 3.2 | 418.8 |
| THA | 5,531 | 2.5 | 71.8 |
| VNM | 11,502 | 5.2 | 182.6 |
| Total | 220,264 | 100.0 | 78.6 |

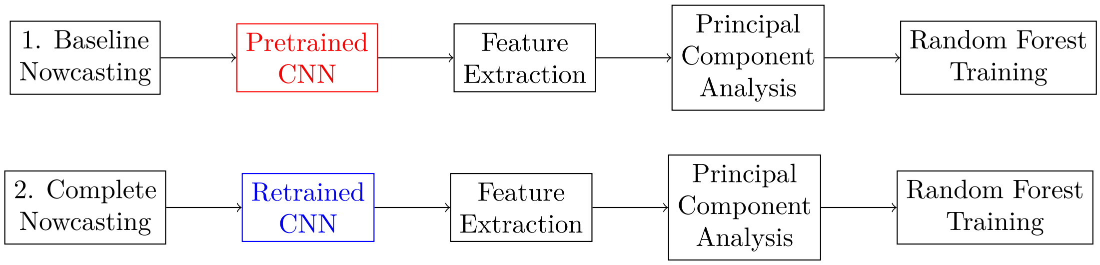
Source: Harmonized Sentinel-2 MSI Level-2A.

## 1.4 Transfer-Learning Models

This section discusses the methodology for predicting GDP per capita, utilizing a sampling frame based on urban areas from the GHS-SMOD dataset and satellite imagery from two distinct sources to forecast label data at the subnational level. The transfer-learning method proves to be valuable in our context where the target data is limited. We propose two transfer learning models to demonstrate how the method can achieve competitive performance in a data-limited environment.

The section begins by establishing a baseline method in which the neural network of the image processor is trained based on the previous study. Subsequently, we introduce a refined approach in which the image feature extractor is retrained using data from 2018 to

Figure 1.3: Transfer-learning process in the two exercises




2019. The retrained network is then integrated into the transfer-learning pipeline, referred to as the complete nowcasting model. This model facilitates the timely assessment of economic trends and developments. Figure 1.3 overviews the steps in the two models, highlighting the main difference between them, which is the use of pretrained or retrained feature extractors.

### 1.4.1 Baseline Nowcasting Implementation

The main approach in this baseline exercise and its subsequent modification is based on the concepts of **nowcasting** and **transfer learning**. Essentially, the nowcasting approach leverages historical data, particularly from 2018 to 2019, to predict an unknown present—the GDP per capita in 2020. This toolbox is particularly valuable for countries that not only lack reliable data but also where regularly published GDP data may experience delays. In critical situations such as natural disasters or civil conflicts, where timely economic information is crucial, nowcasted GDP estimates can serve as a vital resource to fill data gaps. Importantly, this process of training with current data to predict the future can be an ongoing endeavor, with more cases for training continually enhancing both the CNN feature extractor and the RF prediction model.

Transfer learning is a machine learning technique in which knowledge gained from one task is leveraged to improve the performance of a model on a different but related task. Essentially, it enables a model to transfer the skills and insights it has acquired from one domain to another, often reducing the need for extensive training data and computational resources in the new task. In the context of this study, Jean et al. (2016) proposed a transfer learning pipeline. It begins with the original task involving big data. The neural network for image recognition is trained using daytime imagery to predict the nighttime value of the same area, with the underlying assumption that the nighttime value is positively correlated with the economic development of the area. Once the CNN is trained, it is repurposed as the feature extractor to generate important features, as determined by the CNN. These features are then carried over to predict the target task, which has limited data - in our case, it is GDP per capita.

In this exercise, we harness the power of both nowcasting and transfer learning to optimize the prediction of GDP per capita in areas with limited labeled data. The process begins by extracting intricate features from daytime satellite images using **a pre-trained CNN**. To streamline the feature space, PCA is employed to reduce dimensionality while preserving critical information. Subsequently, a Random Forest model is trained, making efficient use of the reduced-feature dataset from previous-period data to predict current GDP per capita. This approach proves especially valuable in situations where labeled data is scarce, enabling accurate predictions. The details are outlined as follows.

### Convolutional Neural Network

A CNN is a type of deep learning model specifically designed for image recognition tasks. It consists of multiple layers that learn to detect various features within images that are important to predict outcome values. It processes images by applying small filters to detect features like edges and textures. These filters slide over the image, and the results pass through an activation function for non-linearity. Pooling layers reduce spatial dimensions, and fully connected layers combine features for predictions. During training, backpropagation adjusts network parameters to minimize prediction errors. CNNs automatically learn and combine features, making them effective for image tasks like classification and object detection.

**Pre-trained Model:** In this baseline exercise, we employ a pre-trained CNN as introduced by Mathur (2020) and Jean et al. (2016) to extract features from sampled Sentinel-2 images for all years in our sample.[7] Mathur (2020)'s CNN training utilizes daytime satellite images from *Planet*, providing higher-resolution data. This model was initially trained to predict nighttime values from nighttime images in countries like Malawi, Nigeria, and Ethiopia. Essentially, it captures the intricate relationship between various daytime surfaces and lighting conditions, a critical insight since nighttime light is strongly associated with economic activity. In the following subsection, we will adapt the transfer learning method by retraining the pre-trained CNN, referred to as the complete nowcasting model, to refine its parameters and acquaint it with lower-resolution images from Sentinel-2.

**Feature Extraction:** Once we gather the pretrained model, we devise the CNN as a feature extractor. We extract 4,096 features that are relevant to predicting nighttime light levels. The size of the extracted features is determined by the choice of the pretrained CNN's architecture, which, in this case, is a Visual Geometry Group model with 11 weight layers (VGG11).

[7] The details and output model are available in Mathur (2020)'s GitHub repository. The purpose of the repository is to replicate Jean et al. (2016)'s training method using more up-to-date tools.

### Principal Component Analysis

PCA is a dimensionality reduction technique used to simplify complex datasets while retaining the most important information. It transforms the original features into a new set of uncorrelated variables called principal components. PCA reduces the dimensionality of the feature space, making it more computationally efficient and less prone to overfitting. It also helps in removing noise and redundancy in the data.

In this exercise, we apply PCA to the 4,096 extracted features, reducing them to a more manageable 10 dimensions while preserving the most significant information.[8] This dimensionality reduction step is crucial for the efficient subsequent training of the Random Forest.

Table 1.5 presents the variances explained by the 10 principal components. Notably, the initial component, referred to as *N0*, accounts for a substantial 22 percent of the total variation in the original set of features. However, as we progress through the subsequent components, the explained variance gradually diminishes. This trend suggests that the first few principal components capture the most critical patterns within the data. Furthermore, the variances tend to stabilize as we move down the list, indicating diminishing returns in terms of explanatory power when including additional components in the analysis.

Table 1.5: Variance explained by 10 principal components, the baseline model

| Component | Variance |
|---|---|
| N0 | 0.22 |
| N1 | 0.10 |
| N2 | 0.06 |
| N3 | 0.04 |
| N4 | 0.03 |
| N5 | 0.03 |
| N6 | 0.03 |
| N7 | 0.02 |
| N8 | 0.02 |
| N9 | 0.02 |

### Random Forest

A Random Forest (RF) model for regression is a machine learning technique designed for predicting numerical values. It operates by forming an ensemble of decision trees that collaborate to enhance prediction accuracy. This ensemble strategy involves

[8] We also attempted dimension reduction to 20 and 100 components but found no substantial improvement in results compared to the 10-dimension model.

training multiple decision trees, with each tree learning from a distinct subset of the dataset. A key characteristic of this model is the introduction of randomness during both data sampling and feature selection processes. This randomness serves to mitigate the risk of overfitting and ensure more robust predictions. The final prediction is generated by aggregating the predictions from all individual trees.

This approach offers several advantages, such as achieving high accuracy, particularly in handling complex and non-linear relationships in data, as well as identifying the most significant features for prediction. However, it's worth noting that the model's results may be less straightforward to interpret, and the computational requirements can be demanding, compared to linear models. Effective tuning of hyperparameters, such as the number of trees and tree depth, is essential for optimizing performance.

In our study, a RF regression is trained using all the data from 2018 to 2019, referred to as the training set. Observations from the year 2020 are used to test the model. Three distinct sets of features are trained in separate models: exclusively nighttime features, exclusively daytime features, and a combination of both nighttime and daytime features. Additionally, we fine-tune the relevant hyperparameters to optimize the performance of the RF model.[9]

It's important to highlight that there can be discrepancies in data publication, particularly in the case of nighttime images, which are typically released annually and with a delay of approximately 6 months after the conclusion of the respective year. In contrast, daytime Sentinel-2 data enjoys more frequent publication, rendering the nowcasting method with daytime features, both the baseline and complete models, more suitable and applicable for the timely prediction of GDP data.

In summary, this transfer learning baseline begins with the pretrained CNN. It is repurposed as a feature extractor, generating 4,096 relevant features from each sampled daytime image. PCA is then applied to reduce the feature dimensionality to 10 while preserving important information. Finally, a RF model is trained on the data from 2018 to 2019, and the best model is used to predict GDP per capita in the 'now' period of 2020. The results are evaluated in the subsequent section.

### 1.4.2 Complete Nowcasting Exercise

In this section, we outline the process of enhancing the baseline toolbox by **retraining the CNN extractor** with Sentinel-2 images from 2018 to 2019 to predict nightlight values as a comparable task to predicting GDP. Given the different sources and resolutions of the pixels used in the pretrained model, concerns may arise regarding the neural

[9] Best parameters from hyperparameter tuning are presented in Tables 1.13 and 1.14 in Appendix A2

network's ability to effectively extract features when presented with Sentinel-2 images of varying scales. This section details the steps taken in this CNN training process.

1. *K-means clustering:* The unsupervised algorithm is employed to classify the daytime images based on their continuous nighttime values. This algorithm categorizes the data into $K$ groups with the aim of minimizing the in-cluster sum of squares. Using this $K$ parameter from the original pre-trained model, we partition the sampled images into three categories based on their brightness during the nighttime. Table 1.6 shows that 58 percent of the images fall into the lowest bin, with VIIRS values ranging from -0.6 to 1.0. The higher-value bins constitute 28 percent and 14 percent of the images, respectively.

Table 1.6: VIIRS Bins

| VIIRS bin | Count | Share | Mean | Min | Max |
|---|---|---|---|---|---|
| 0 | 239,183 | 0.58 | 0.27 | -0.60 | 1.07 |
| 1 | 115,609 | 0.28 | 2.69 | 1.07 | 6.62 |
| 2 | 60,097 | 0.14 | 19.30 | 6.62 | 2392.02 |

2. *Data Splitting:* To establish a balanced training dataset, we randomly allocate 80 percent of the remaining non-transparent images from each country to the training group, with the remaining 20 percent assigned to the validation group. This approach ensures a consistent training-to-validation ratio across countries, guarding against potential biases that may arise from the influence of any single country.

3. *Fine-tuning the CNN Model:* The retraining process entails fine-tuning all layers of the CNN model over 20 epochs. Notably, the R-squared metrics remain stable around 0.84 to 0.85 from the 16th epoch onwards, with the highest validation R-squared reaching 0.85. This illustrates the effectiveness of the retraining process in capturing valuable features from the Sentinel-2 images.

**Feature Extraction and RF Training:** We repurpose the retrained CNN into a new feature extractor and commence the feature extraction process. We apply PCA to further enhance the extracted features, and their explained variances are listed in Table 1.11. Additionally, we carry out RF training, both with and without nighttime features, using the same training and testing groups established in the baseline exercise. This consistent data partitioning enables a meaningful evaluation of the results.

## 1.5 Results

This section presents results from all the toolboxes detailed in the previous section and those to be compared with the benchmark. We begin by discussing the performance of the proposed methods. Furthermore, we explore daytime features that are deemed important by the transfer-learning method. Afterward, we present showcase outcomes for countries within the sample. The results indicate that the overall performances are influenced by the Chinese sample, which slightly underperforms compared to the other countries. Finally, the method is deployed to predict the change in GDP per capita, and it is realized that the performance is subpar compared to the previous study.

### 1.5.1 Performances

To assess performances of the methods, Table 1.7 presents the R-squared values of the baseline and nowcasting models using different sets of features: only nighttime (*ntl*), only daytime (*sentinel*), and both (*ntl_sentinel*). As mentioned earlier, the training set comprises data from 2018 and 2019, while the test set consists of data from 2020. To evaluate the performance, we consider the 2020 $R^2$ values, which represent out-of-sample predictions.

In the baseline method, where the CNN extractor is pre-trained using the previous study, the prediction performance improves when both daytime and nighttime features are combined to predict the labels. The test R-squared increases from 0.63 or 0.65 to 0.76. It's worth noting that in models where only nighttime or daytime features are used as predictors, the nighttime model slightly outperforms the daytime model in explaining GDP per capita. This result remains robust when the evaluation metric is changed to mean square errors, as shown in Table 1.12 in Appendix A1.

The complete nowcasting model, in which the CNN extractor is trained and fine-tuned using Sentinel-2 images, improves the out-of-sample prediction when only daytime features are used as predictors. The 2020 $R^2$ increases from 0.63 in the baseline to 0.70 in the complete nowcasting model. However, there is only a slight gain in goodness of fit when the full set of features is deployed. Its R-squared increases slightly from 0.76 in the baseline to 0.77. The mean squared errors in Table 1.12 also reflect the same pattern.

Figure 1.4 displays the out-of-sample predicted and actual GDP per capita from the complete model, *ntl_sentinel_best_nowcast*. It is notable that the majority of underestimation occurs at high values, a pattern similar to that observed in other studies predicting economic outcomes from remote-sensing data.[10] This pattern of underpredicted at high-

[10]Figure 1.12, featured in a later subsection, reaffirms the pattern of overestimation among low-income and underestimation among high-income subregions across countries in the sample.

economic-outcome regions and overpredicted at low-economic-outcome regions can be seen from the previous studies.[11]

We further explore what determines the deviation from the "truth" or where the model performs poorly. We address this question by conducting linear Lasso regression on error terms, defined as $GDPpc_i - \hat{GDPpc_i}$, in which a positive error indicates underestimation. Using all the features discussed in the study, the Lasso regression selects significant predictors that explain the error. Figure 1.5 illustrates the size and direction of non-zero coefficients. The results confirm underestimation among high-income, high *N0*, and brighter subregions (high *total_ntl*).

One observation to highlight from the Lasso regression is that the sizes of subregions and urban areas exhibit weak influences on the deviation. This implies that the method performs well across areas of different sizes. Similarly, population size has no influence on the estimated error, as it is eliminated by the regression. As a result, one can infer the external validity of the method, indicating that the model performs competitively regardless of the subregion's size and population size.

To evaluate the effectiveness of the proposed transfer learning methods in the context of developing countries, we shift the predicted target from GDP per capita to subregional GDP values and compare their performances with the benchmark study, which reports an R-squared range of 0.82 to 0.90. Utilizing our complete nowcasting model, we find that the performance is 0.74 when only daytime features are used, while the full-feature model achieves an impressive 0.94 (see Rows 6 and 7 in Table 1.7).[12] These results provide reassurance that the transfer learning models perform competitively when compared to directly trained CNN models in data-rich contexts.

Beyond performance, feasibility is another crucial aspect to consider when comparing with the benchmark study, particularly in data-limited environments where ground truth is scarce. Training a deep neural network, such as a CNN, to directly predict economic value requires that the raw data structure be compatible with the CNN architecture. This typically necessitates input in the form of images adhering to specific standards. The benchmark study effectively employed the CNN method due to the availability of high-resolution, labeled data at the block-group level, which was derived from the distribution of population data at a finer block level. This approach enabled the interpolation of block-level income data into image-level representations, thereby ensuring compatibility with CNNs.

However, in our context, where labels are available only at the subregion level, in-

[11]For example, Jean et al. (2016), ADB (2020),

[12]See Figure 1.16 in Appendix A2 for the scatter plot between the predicted and actual values.

terpolating these labels to image-level data suitable for CNN input would require making significant assumptions and acquiring additional finer-resolution datasets. These requirements render the implementation of CNNs in a limited-data setting nearly impossible. Furthermore, the additional methods needed for such interpolation might raise concerns about validity and could potentially compromise the model's performance.

In contrast, the proposed transfer learning method offers a more feasible approach. It provides a consistent template for various types of subregions with differing structures, thereby maintaining external validity across different contexts.

Table 1.7: R-squared: baseline and nowcasting models

| Row | Model | Label | Train $R^2$ | 2020 $R^2$ | Total $R^2$ |
|---|---|---|---|---|---|
| 1 | ntl_best_baseline | GDP per capita | 0.76 | 0.65 | 0.73 |
| 2 | sentinel_best_baseline | GDP per capita | 0.94 | 0.63 | 0.84 |
| 3 | ntl_sentinel_best_baseline | GDP per capita | 0.96 | 0.76 | 0.90 |
| 4 | sentinel_best_nowcast | GDP per capita | 0.93 | 0.70 | 0.86 |
| 5 | ntl_sentinel_best_nowcast | GDP per capita | 0.96 | 0.77 | 0.90 |
| 6 | sentinel_best_nowcast | GDP | 0.96 | 0.74 | 0.88 |
| 7 | ntl_sentinel_best_nowcast | GDP | 0.98 | 0.94 | 0.97 |

Notes: Rows 1-3 are related to the baseline model; Rows 4-5 pertain to the complete nowcasting model; Rows 6-7 represent the outcomes derived from the complete nowcasting model for GDP prediction. The training group contains data of 2018 and 2019 while the test group is the sample of 2020.

Figure 1.4: Actual vs. predicted GDP per capita, *ntl_sentinel_best_nowcast*

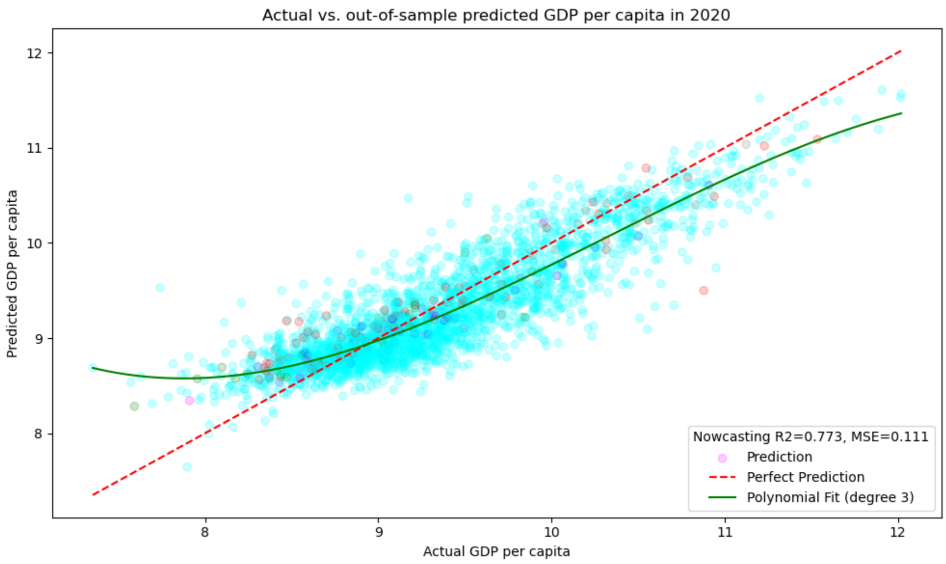
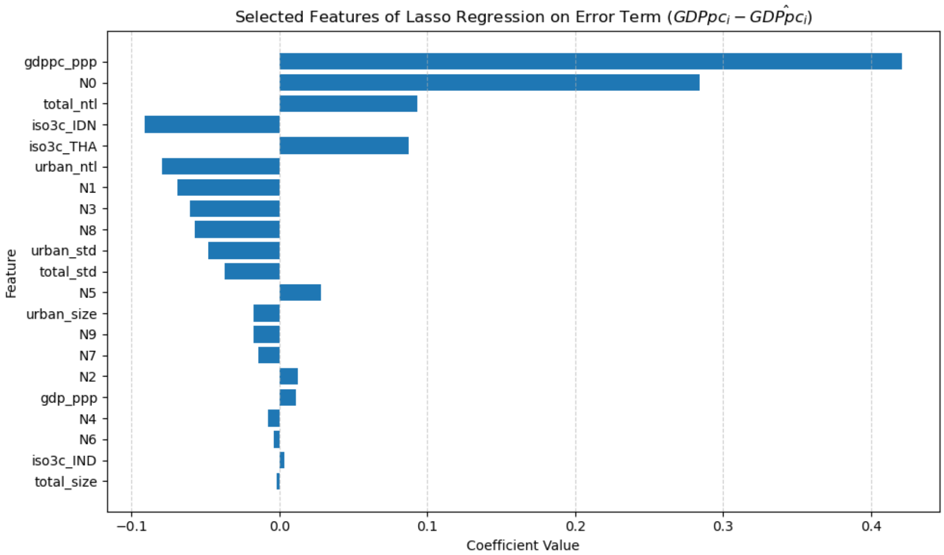


Figure 1.5: Features selected by Lasso regression on the error term, *ntl_sentinel_best_nowcast*




Note: The base category of the country variables is Malaysia, where the subgroup's performance is the highest.

### 1.5.2 Important Features

One caveat of the RF regression is its inability to interpret the relationships between predictors and the target outcome. This limitation arises from its ensemble characteristic, where a random subset of features is selected in each iteration to build a decision tree, resulting in a "random" forest with many different trees, each using a different set of sampled features.

However, the method does provide a measure of feature importance by calculating the accuracy loss when a feature is excluded. This subsection first discusses the importance of features in the complete nowcasting model, as computed on the test dataset. It then further illustrates the relationships between important daytime features and GDP per capita. Utilizing daytime satellite images from Sentinel-2 with extreme values of key features, the subsection further demonstrates the types of global surfaces that contribute the most and the least to the subregional economic development.

Figure 1.6 compares the importance of all features in the model and establishes that urban brightness and the first principal component ($N0$) are the most crucial predictors of GDP per capita. Other daytime components that have made it to the top-10 list are $N3$, $N6$, and $N2$, respectively. Meanwhile, all nighttime features are also in the top-10 list. The subregion brightness (*total_ntl*) ranks as the third most important feature, but its significance is substantially lower than that of the top two. This is also true for the rest of the features, as they contribute limited accuracy to the model.[13]

[13] Figures 1.17 and 1.18 in Appendix A2 display important features from the model with only nighttime and only daytime features, respectively. They also reinforce the same conclusion that urban brightness and the first component are the most important in their respective categories, significantly surpassing the runner-ups.

Further investigation of these important daytime features involves plotting their behavior against GDP per capita in Figure 1.7. These features exhibit non-linear behavior, with the most significant variation observed when contrasting $N0$ against the outcome in the top panel, as expected. On the other hand, less crucial daytime features exhibit a more flattened relationship with the outcome, or, at the very least, the majority of datapoints are situated in flatter regions, as compared to $N0$.

To address the question of how the global surface appears in an area that significantly contributes to income, we examine daytime satellite images with extreme values of these important predictors in Figure 1.8. In the top two panels, the daytime images with the highest and lowest values of $N0$ reveal that high-$N0$ images, representing low income per capita, depict green areas with small clusters of man-made structures, while low-$N0$ images, indicating high-income areas, contain well-planned urban areas with clusters of high-rise buildings.

The presentation of satellite images provides a more meticulous view of how the CNN has learned to distinguish between low- and high-income areas. For example, images with low values of $N3$ (indicative of high income) and $N6$ (indicative of low income) both cover residential areas but in two different countries, Indonesia in the *Low N3* panel and India in the *Low N6* panel. Both sets of images depict small housing units on flat terrain with scattered green spaces. One striking visible contrast between the two is the color of the rooftops on these buildings. In Indonesia (Java and Bali), the rooftops are brown or red, resembling terracotta-clay tiles used for roofing. In the Indian states, the rooftop color is white-grey, typically made of concrete or cement, which are more affordable materials. This observation highlights the impressive capacity of the neural network to materialize its knowledge by learning basic shapes and colors, and mapping these learned patterns to nighttime light.

Figure 1.6: Top 10 most important features, *ntl_sentinel_best_nowcast*

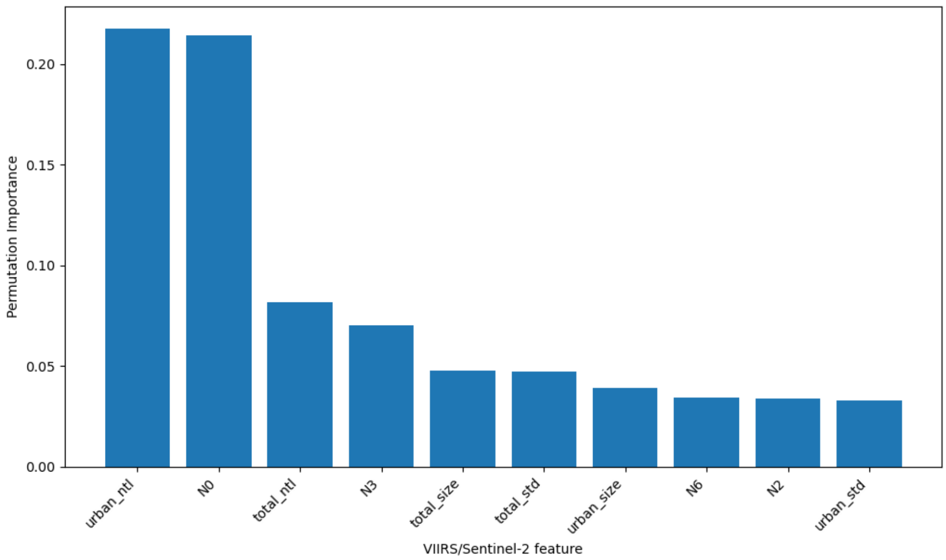


Figure 1.7: GDP per capita and Sentinel-2 features, *ntl_sentinel_best_nowcast*

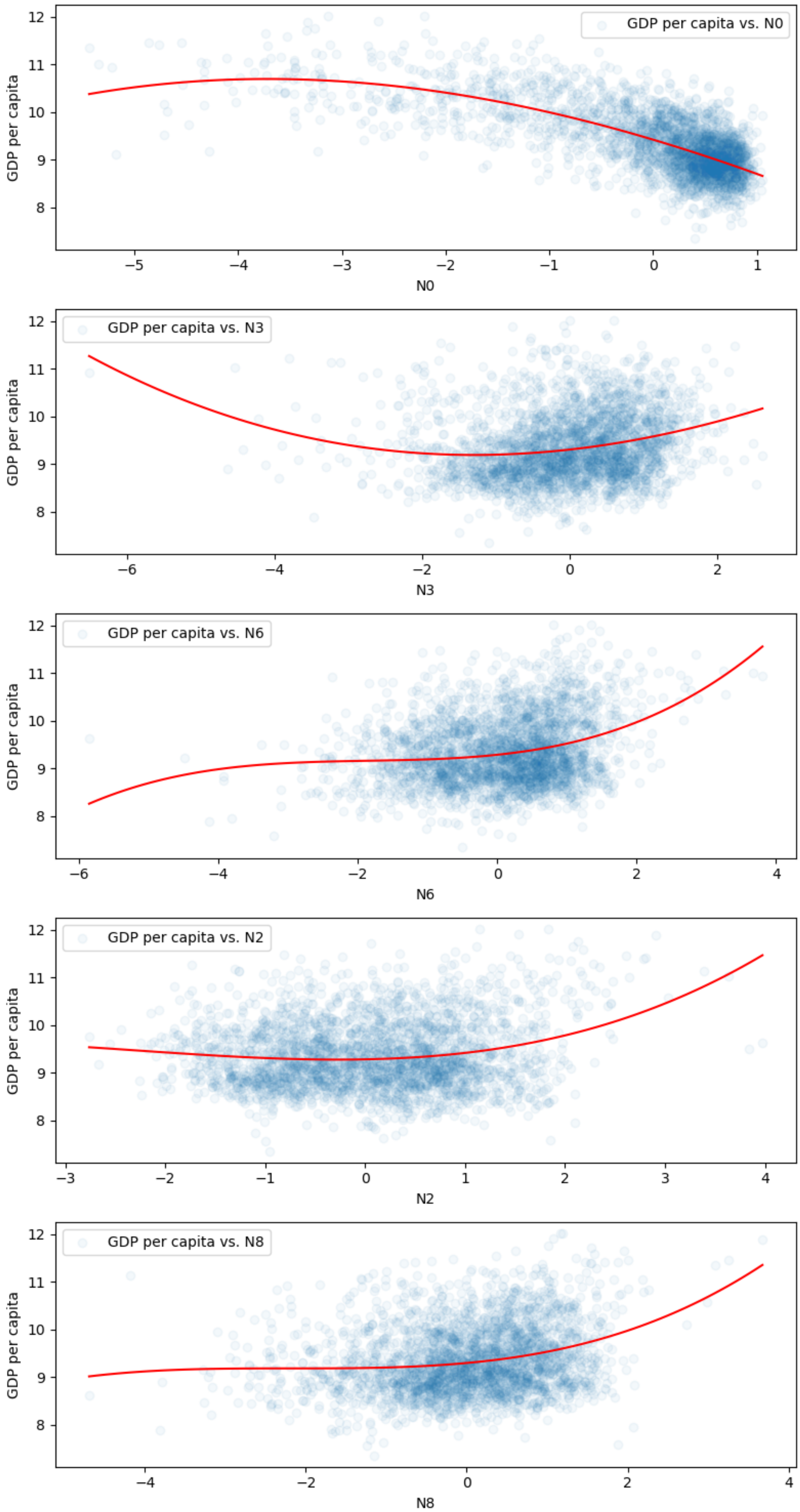
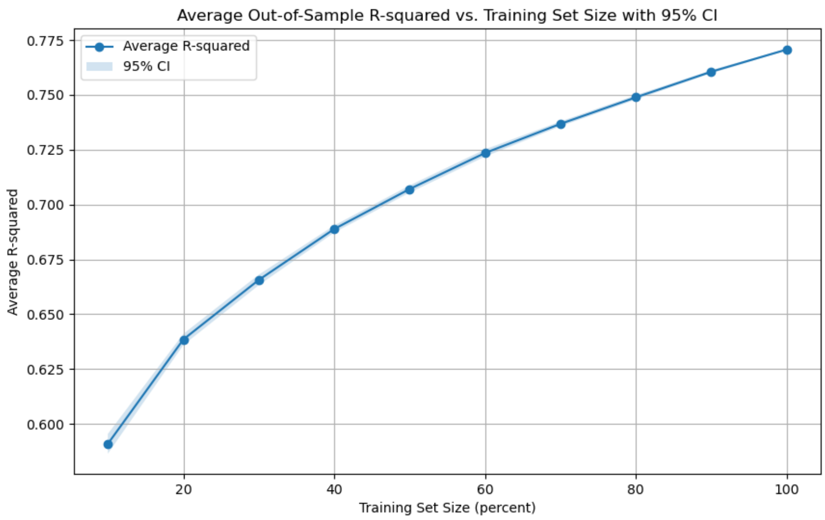


Figure 1.8: Images with high and low daytime features, *ntl_sentinel_best_nowcast*

### 1.5.3 Varying Training-Set Sizes

This subsection discusses a robustness test involving varying training sample sizes to assess the stability of the results and how accuracy improves as the training set grows larger. Figure 1.9 illustrates the average R-squared values by training set size, where the 100-percent size corresponds to the result in Row 5 of Table 1.7. Within each sample size, we bootstrap the training set with the corresponding size 50 times and train RF regression using the subsample to predict the outcome variable in 2020. The average performance, in terms of R-squared, has significantly improved when we add one more year of labels, which constitutes 50 percent of the total training set. R-squared has increased from 0.71 at the 50-percent training size to 0.77 at the full training size. It's worth noting that as time progresses and new labels from another year are included in the training, we expect the performance to further improve to around 0.79 to 0.80, and the training may reach a plateau with a couple more years of data.[14]

Figure 1.9: R-squared by size of training set, *ntl_sentinel_best_nowcast*




### 1.5.4 Results by Country

One concern that may arise from this exercise is that the majority of the subregions are Chinese counties, accounting for almost 94 percent of the test sample. It's possible that the satisfactory performance mentioned above is primarily driven by the Chinese sample, while the samples from other countries perform poorly. This subsection dissects the aggregate results by country to ensure that the predictions in the other countries are a better fit than the overall result and those specific to China.

Figures 1.10 and 1.11 contain R-squared values and mean squared errors for all countries in the test group. In both metrics, they indicate that the overall performances

[14] In Appendix A2, Figure 1.19 plots the average mean squared errors for the same test.

are close to those of China. However, samples from all the countries perform better than the Chinese subgroup, except for India, where its R-squared is lower than China's but with a lower MSE. This means that the overall result is indeed influenced by the less well-fitted sample from China, but the transfer-learning toolbox generally performs well in the other countries in the sample. For example, the complete nowcasting method can predict the GDP per capita of Malaysian subregions with impressive accuracy, as evidenced by the highest R-squared value of 0.89 and lowest mean squared error of 0.03. Meanwhile, the goodness of fit for India's prediction is the lowest at 0.67, but the mean square error is lower than that of China and Thailand at 0.09.

The underperformance in predicting Chinese GDP may be attributed to data quality issues, a speculation in line with its lowest Statistical Capacity Score among the sampled countries, as reported by the World Bank (Cameron et al., 2019). Chen et al. (2019) proposes a plausible explanation for this discrepancy, pointing to the motivations of local governments that prioritize meeting GDP and investment performance targets. Despite statistics being compiled by local branches of China's National Bureau of Statistics (NBS), the input data are less detailed compared to the national GDP computation carried out by the central NBS office. In practice, local NBS offices receive funding from local governments, and the evaluation and promotion of officials within local NBS are influenced by local government priorities. This complex relationship between local governance and statistical practices contributes to the observed variations in Chinese GDP data.

To further investigate the scatter plot of the predicted values and the actual GDP per capita, Figure 1.12 splits the overall sample by country and shows how the predictions behave across countries. The India sample performs poorly in the low-income subregions, where the model overestimates the target to a significant extent. In general, most of the countries, except the Philippines, fit into the aggregate pattern in which the model overstates GDP per capita in the low-income regions while it underpredicts in the high-income areas.

Figure 1.10: R-squared by country, *ntl_sentinel_best_nowcast*

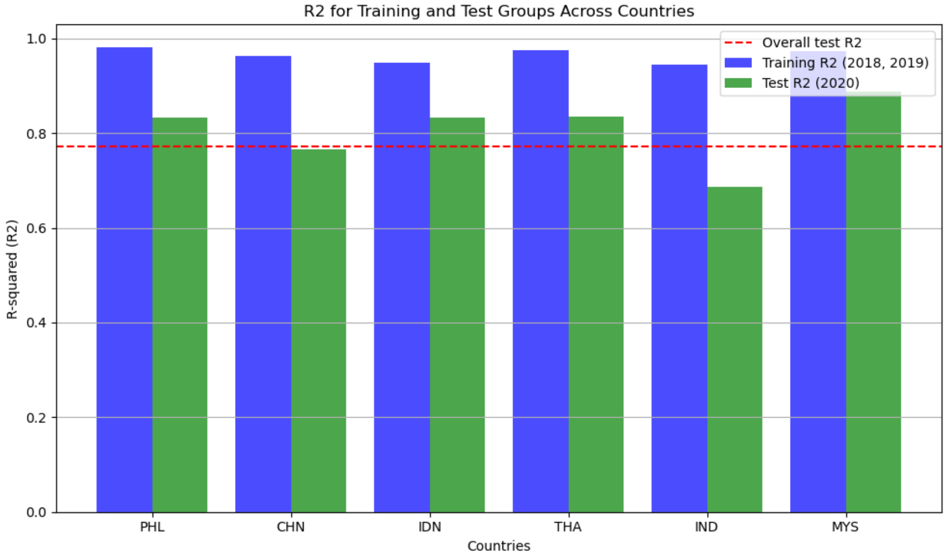



Figure 1.11: Mean squared error by country, *ntl_sentinel_best_nowcast*

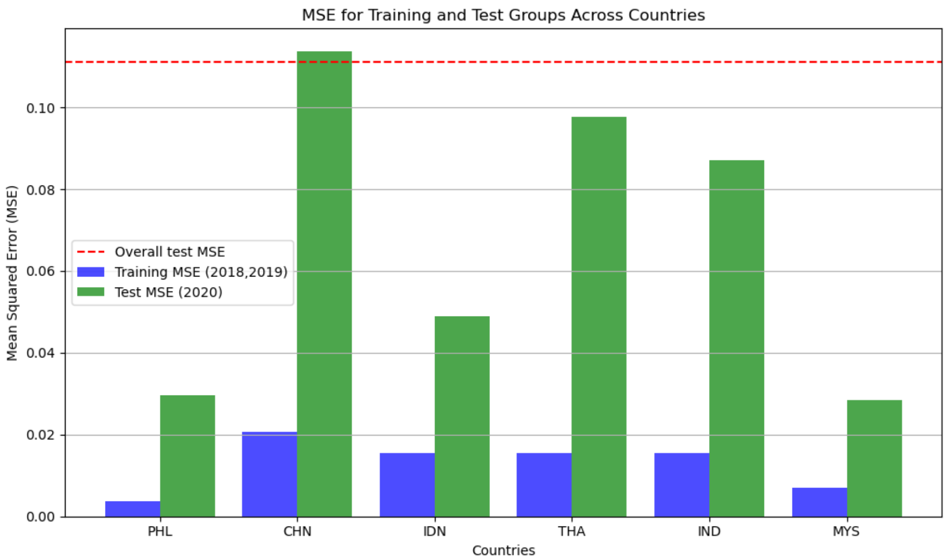
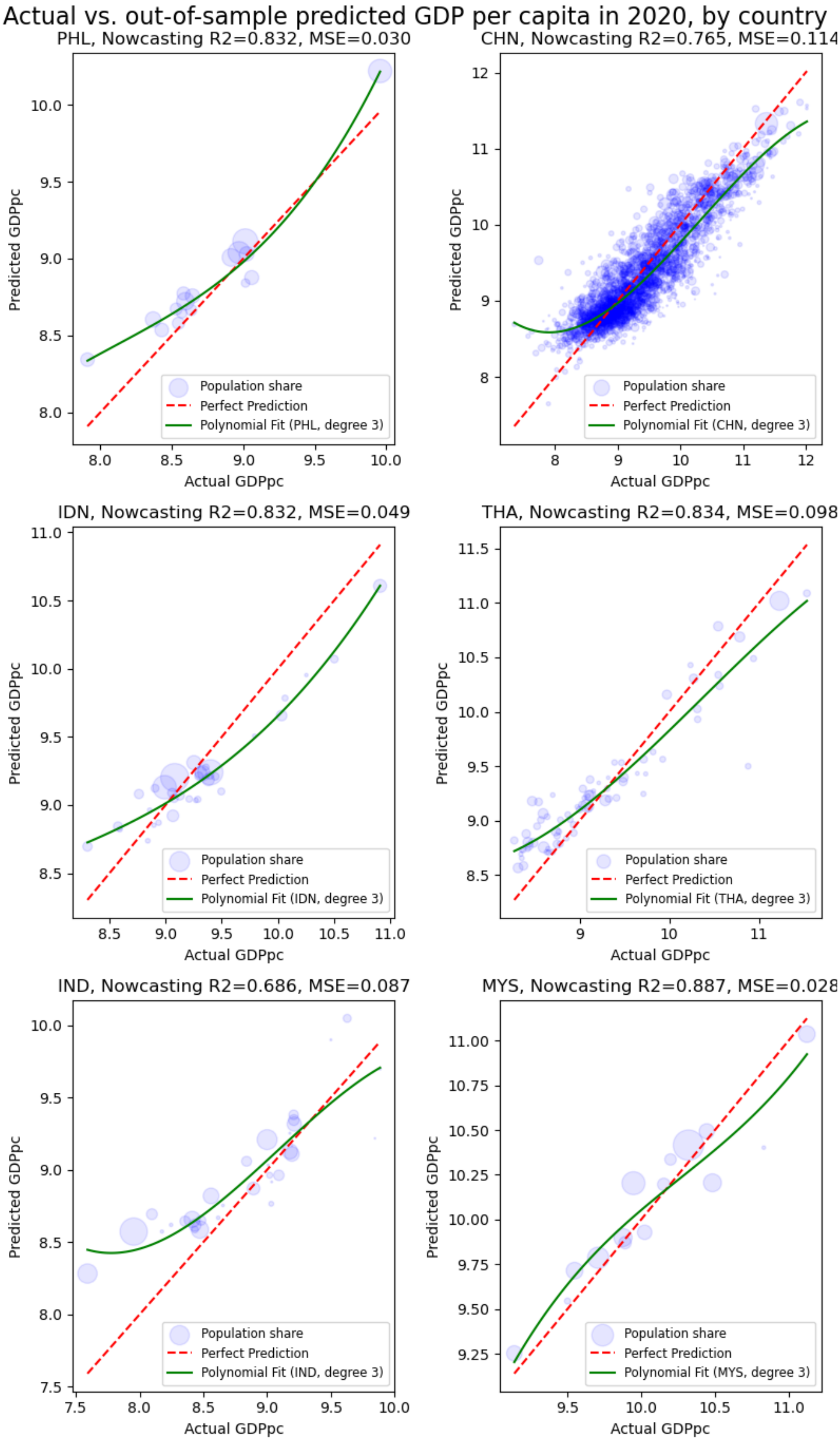


Figure 1.12: Actual vs. predicted GDP per capita by country, *ntl_sentinel_best_nowcast*



### 1.5.5 Predicting Changes

A more advanced application of the model involves predicting GDP per capita growth rates, particularly during challenging periods such as the 2020 pandemic that impacted many countries. Building on the previously retrained CNN, we propose three scenarios for undertaking this task, each depending on the data environment and the train-test split. Table 1.8 outlines the specifics of each scenario, highlighting their differences in terms of data split, predictors, and the appropriate data environment for their application.

The first scenario extends the previous task of estimating GDP per capita by further computing the GDP per capita growth rates from the predicted values. The *nowcasting method* used in this scenario emphasizes that the split between training and test groups ensures that all information from the previous period is utilized to predict current growth rates. In this data split, the training set consists of observations from 2019, while the test set comprises samples from 2020. The predictors used in the Random Forest (RF) training are satellite features, including both nighttime and daytime imagery. The RF model is first applied to predict GDP per capita, after which the growth rates are computed from the predicted values and compared against the actual growth rates. This method is straightforward and minimally demanding, making it suitable for data environments where official subregional GDP and population data are not available.

*Scenario 2* is proposed under the assumption that official GDP and population data are available at the subregional level. Consequently, these additional features from the previous period are incorporated into the RF model, which is then trained to predict growth rates directly. This scenario is applicable only when official data from the preceding period is accessible to inform the model's predictions.

The final scenario departs from the nowcasting approach by splitting the train-test dataset cross-sectionally, with a 70/30 ratio. This method ensures that the entire time series of a subregion in the training set is trained during both economic expansions and downturns. The same set of features as in the previous scenario is employed in the RF model, making this approach viable only in data environments where official subregional data are available.

Table 1.8: Three scenarios for predicting growth rates

| Scenario | Data Split | Predictors | RF Label | Data Environment |
|---|---|---|---|---|
| *1* | *Nowcasting method* | NTL and daytime features from previous and current periods | GDP per capita, then calculate the growth rates of the predicted values. | Subregions without GDP and population data collection |
| *2* | *Nowcasting method* | NTL and daytime features from previous and current periods, and subnational GDP and population from the previous period | Growth rates | Subregions with official data |
| *3* | *Cross-sectional split* | NTL and daytime features from previous and current periods, and subnational GDP and population from the previous period | Growth rates | Subregions with official data |

Notes: *Nowcasting method* means the training group consists of data from 2019, while the test group comprises data from 2020. The *Cross-sectional split* divides the sample cross-sectionally, ensuring a balanced sample from all countries.

The overall performance results are presented in Table 1.9 in terms of R-squared. The first two scenarios, which utilize the nowcasting approach, yield negative R-squared values for the test group, specifically -1.46 and -0.08, respectively. These results indicate that the methods do not perform well, even when the RF model is trained to predict growth rates directly in *Scenario 2*.[15]

However, when the data split is conducted cross-sectionally, with an increase in the training size from 50/50 to 70/30, the R-squared improves to 0.2 in *Scenario 3*. The scatter plot in Figure 1.13 shows that the predicted values align well with actual data for small to moderate positive growth rates but tend to overestimate negative growth rates and underestimate extremely high growth rates.

Out-of-sample performance varies across countries, as illustrated in Figure 1.14, where an upward bias is generally observed in nearly every country. Figure 1.24 in Appendix A2.2 provides further details, indicating that predictions are most accurate for India, China, and Indonesia, while they perform worst for Thailand and the Philippines.

We further investigate the factors that most influence estimation errors. Figure 1.15 presents selected factors from the Lasso regression, confirming that underestimation is prevalent in subregions with high economic growth, while overestimation is common in subregions experiencing significant economic downturns, due to the high coefficient

[15] See scatter plots and performance across countries in Appendix A2.2.

of *gdppc_ppp_d*. The second most important factor is population growth (*pop_d*), with subregions experiencing high population growth tending to face overestimation of GDP per capita growth rates. These findings highlight the limitations of this method in accurately estimating growth rates during economic crises, particularly when migration further exacerbates estimation deviations.

To compare the performance with the benchmark study, we adjusted the predicted target to the growth rates of GDP values. The results, presented in Row 4 of Table 1.9, show that even the best-performing method, *Scenario 3*, underperforms relative to the benchmark method, where R-squared values range between 0.3 and 0.4.

Two potential reasons for this underperformance are identified. First, the retrained CNN used to generate important daytime features was trained to predict nighttime values, not changes in those values. Consequently, the knowledge gained from this previous task might be less applicable to the new task of predicting GDP growth in the RF model. A significant revision of the neural network to better handle time-series data might improve the results in this area. Secondly, the sample size in the RF model training is significantly smaller than in the previous exercise of predicting GDP values. It is important to note that when constructing the panel data with the target value of GDP growth rate, the data points from the first period of 2018 are excluded due to the lack of the target label, which requires GDP data from the prior year. With one-third of the data points missing, the RF model's training may suffer from reduced performance. However, as time progresses and more data points become available, this concern may be alleviated. This chapter serves as a foundational starting point, with potential improvements possible through refining the deep learning model and incorporating additional years of data into the training process.

Table 1.9: R-squared: Growth predictions

| Row | Model | Label | Train $R^2$ | Test $R^2$ | Total $R^2$ |
|---|---|---|---|---|---|
| 1 | Scenario 1 | GDPpc Growth | nan | -1.46 | nan |
| 2 | Scenario 2 | GDPpc Growth | 0.72 | -0.08 | 0.39 |
| 3 | Scenario 3 | GDPpc Growth | 0.59 | 0.20 | 0.48 |
| 4 | Scenario 3 | GDP Growth | 0.55 | 0.17 | 0.45 |

Notes: In Scenarios 1 and 2, the training group consists of data from 2019, while the test group comprises data from 2020. In Scenario 3, the training and test groups are split cross-sectionally, ensuring a balanced sample from all countries.

Figure 1.13: Actual vs. predicted growth rate in GDP per capita, *Scenario 3*

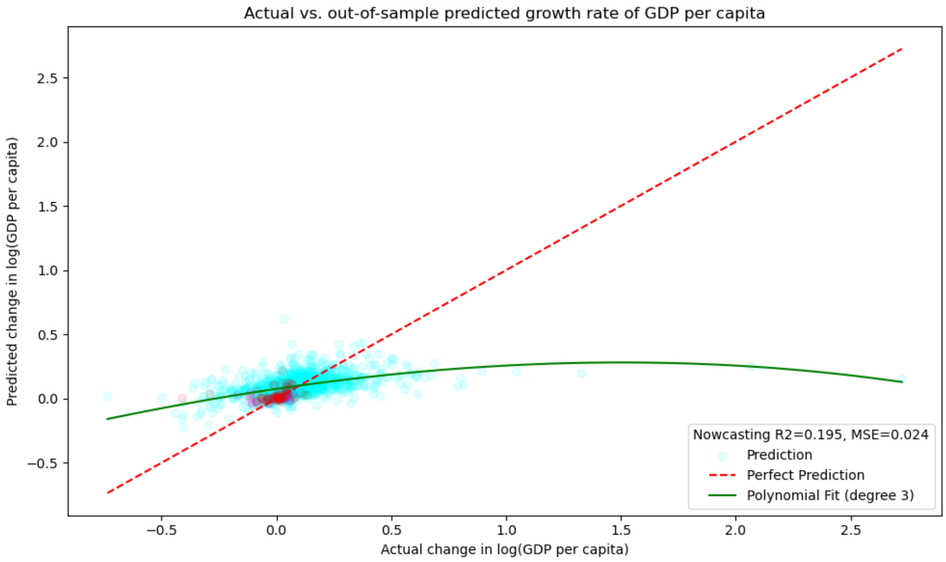



Figure 1.14: Average actual and predicted growth rate with 95% confidence intervals, *Scenario 3*

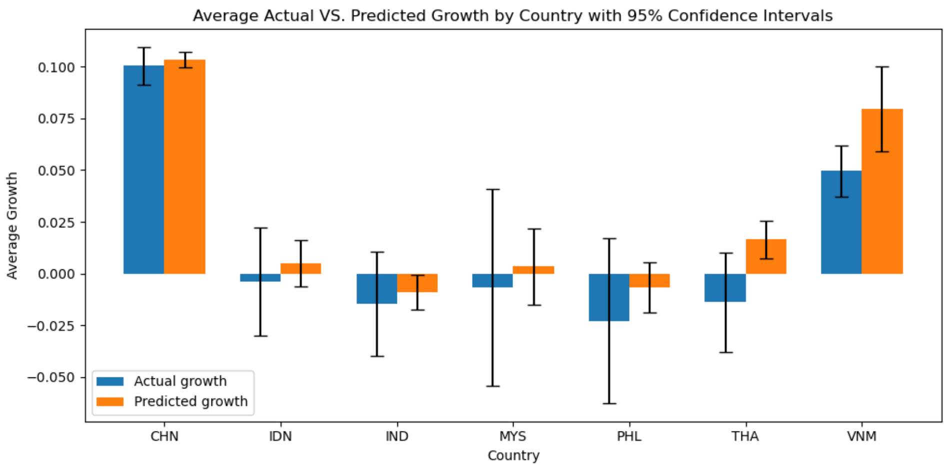
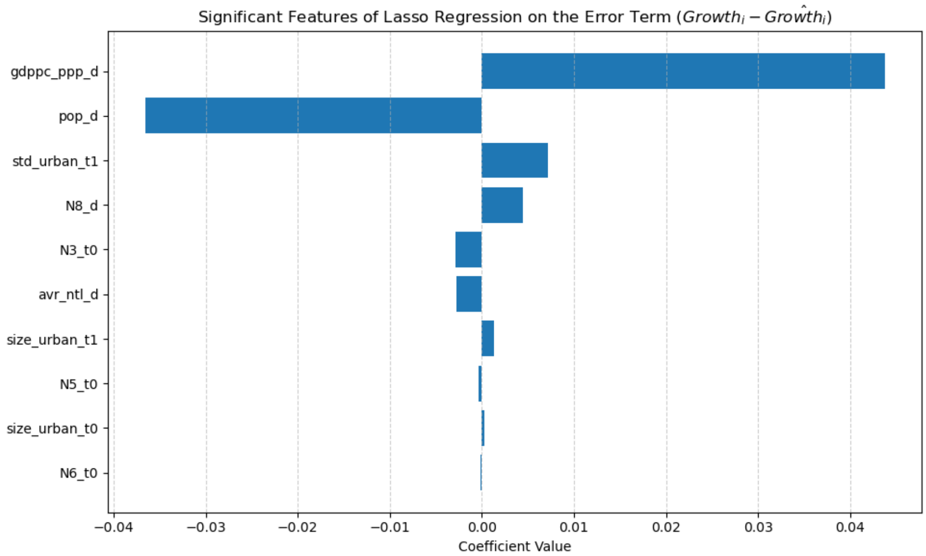



Figure 1.15: Features selected by Lasso regression on the error term, *Scenario 3*



## 1.6 Summary

Subnational GDP data is often lacking, particularly in developing countries, which poses challenges for policymakers and hinders economic development efforts. To address this issue and enhance data reliability, remote-sensing data has become a valuable resource, offering publicly available databases that have improved with advancements in computer science.

The main focus of this chapter is Southeast Asia, a region known for its diverse economic conditions and limited data availability. Using satellite images and machine learning, the study introduces a toolbox for estimating subnational GDP per capita. India and China are included in the study to increase the sample size.

This chapter contributes to the existing literature by introducing a methodology for estimating subnational economic indicators using satellite imagery and machine learning techniques. The ultimate goal is to address data gaps and facilitate economic research in regions with limited data availability. In consideration of implementation costs and data availability within the context of developing countries, we leverage modest databases and the transfer learning method to construct a toolbox for estimating GDP per capita. This approach proves effective in achieving competitive performance.

The proposed toolbox consists of two transfer learning exercises. The first exercise involves a baseline model that utilizes a pretrained neural network from a previous study to extract image features from the Sentinel-2 daytime dataset. Principal component analysis is then applied to reduce the dimension of the extracted daytime features to 10, making the data more manageable. Using three sets of predictors (daytime, nighttime VIIRS, and both daytime and nighttime features), we train a Random Forest regression model to predict GDP per capita. The training set includes data from 2018 and 2019, and the test set consists of data from 2020, a method commonly referred to as nowcasting.

The second transfer learning exercise involves complete training, where the only difference from the first exercise is that the neural network is fine-tuned using imagery from 2018 and 2019. This retrained feature extractor significantly improves performance, especially when the Random Forest regression model is trained using only daytime features. As a result, the complete transfer-learning model with the full set of features achieves high performance, with an R-squared of 0.77 or an MSE of 0.11 when predicting out-of-sample GDP per capita.

Furthermore, the Lasso regression analysis indicates that the error terms associated with the predicted values are rarely influenced by the area size and are not correlated with population size. This finding supports the external validity of the method, suggesting it can be applied to other subregions with varying sizes and populations.

Although the benchmark exercise, which uses GDP as the target variable, performs well compared to previous studies, the proposed method faces challenges in accurately predicting changes or growth in GDP per capita, highlighting certain limitations. It is believed that improving performance may require revising the neural network training specifically for the task of predicting growth rates and increasing the number of data points in the Random Forest training.

One practical issue to consider is the timing of predictions. The full set of features can be curated six months after the end of a year due to the delay in nighttime VIIRS data publication. On the other hand, daytime features from Sentinel-2 are available within three weeks after the relevant period, and the performance of the model that solely uses the daytime data is decent, with an R-squared value of 0.70. This fast but less accurate method can also be an option for policymakers in developing countries, especially during crises.

# A1 Tables

## A1.1 Data

Table 1.10: Descriptive statistics of nighttime light features 2019

| ISO3c | total_std | | | | urban_std | | | | urban_size | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | mean | std | min | max | mean | std | min | max | mean | std | min | max |
| CHN | 1.20 | 0.72 | 0.01 | 4.60 | 1.61 | 0.64 | 0.00 | 4.36 | 3.85 | 0.93 | 0.69 | 5.98 |
| IDN | 1.07 | 0.58 | 0.25 | 2.64 | 1.66 | 0.48 | 0.82 | 2.98 | 5.51 | 0.78 | 3.47 | 6.22 |
| IND | 1.20 | 0.56 | 0.24 | 2.99 | 1.58 | 0.52 | 0.69 | 2.91 | 5.71 | 0.89 | 3.26 | 6.22 |
| MYS | 2.30 | 0.82 | 1.02 | 3.85 | 2.82 | 0.79 | 2.00 | 5.23 | 4.81 | 0.87 | 2.71 | 6.03 |
| PHL | 0.85 | 0.63 | 0.23 | 2.91 | 1.22 | 0.56 | 0.41 | 2.85 | 5.97 | 0.42 | 4.85 | 6.22 |
| THA | 1.20 | 0.55 | 0.30 | 2.92 | 1.85 | 0.38 | 1.15 | 3.25 | 3.99 | 0.78 | 1.95 | 5.67 |
| VNM | 1.28 | 0.84 | 0.18 | 4.27 | 1.67 | 0.78 | 0.59 | 4.72 | 5.02 | 0.69 | 3.22 | 6.22 |

Source: Author's tabulation from VIIRS.
Notes: *total_std* denotes the standard deviation of the VIIRs values in a region, while *urban_std* represents the standard deviation of the urban area. *urban_size* corresponds to the subregion's urban size. Note that all features undergo a standardization for the analysis.

Table 1.11: Variance explained by 10 principal components, the complete nowcasting model

| Component | Variance |
|---|---|
| N0 | 0.51 |
| N1 | 0.05 |
| N2 | 0.04 |
| N3 | 0.03 |
| N4 | 0.03 |
| N5 | 0.02 |
| N6 | 0.02 |
| N7 | 0.01 |
| N8 | 0.01 |
| N9 | 0.01 |

## A1.2 Predicting GDP per Capita

Table 1.12: Mean-squared error: baseline and nowcasting models

| Row | Model | Label | Train MSE | 2020 MSE | Total MSE |
|---|---|---|---|---|---|
| 1 | ntl_best_baseline | GDP per capita | 0.1275 | 0.1696 | 0.1417 |
| 2 | sentinel_best_baseline | GDP per capita | 0.0308 | 0.1806 | 0.0815 |
| 3 | ntl_sentinel_best_baseline | GDP per capita | 0.0215 | 0.1156 | 0.0534 |
| 4 | sentinel_best_nowcast | GDP per capita | 0.0362 | 0.1479 | 0.0740 |
| 5 | ntl_sentinel_best_nowcast | GDP per capita | 0.0198 | 0.1110 | 0.0507 |
| 6 | sentinel_best_nowcast | GDP | 0.0718 | 0.4078 | 0.1855 |
| 7 | ntl_sentinel_best_nowcast | GDP | 0.0258 | 0.0974 | 0.0500 |

Notes: Rows 1-3 are related to the baseline model; Rows 4-5 pertain to the complete nowcasting model; Rows 6-7 represent the outcomes derived from the complete nowcasting model for GDP prediction. The training group contains data from 2018 to 2019.

Table 1.13: Best parameters: baseline and nowcasting models

| Row | Model | Label | ccp_alpha | max_depth | max_features | max_leaf_nodes | max_samples | min_impurity_decrease |
|---|---|---|---|---|---|---|---|---|
| 1 | ntl_best_baseline | GDP per capita | 0 | 10 | log2 | 0 | 0 | 0 |
| 2 | sentinel_best_baseline | GDP per capita | 0 | 0 | 5 | 0 | 0 | 0 |
| 3 | ntl_sentinel_best_baseline | GDP per capita | 0 | 0 | 5 | 0 | 0 | 0 |
| 4 | sentinel_best_nowcast | GDP per capita | 0 | 0 | 5 | 0 | 0 | 0 |
| 5 | ntl_sentinel_best_nowcast | GDP per capita | 0 | 0 | 5 | 0 | 0 | 0 |
| 6 | sentinel_best_nowcast | GDP | 0 | 0 | log2 | 0 | 0 | 0 |
| 7 | ntl_sentinel_best_nowcast | GDP | 0 | 0 | 5 | 0 | 0 | 0 |

Table 1.14: Best parameters: baseline and nowcasting models (Continue)

| Row | Model | Label | min_samples_leaf | min_samples_split | min_weight_fraction_leaf | n_estimators | n_jobs | warm_start |
|---|---|---|---|---|---|---|---|---|
| 1 | ntl_best_baseline | GDP per capita | 2 | 8 | 0 | 200 | 0 | False |
| 2 | sentinel_best_baseline | GDP per capita | 1 | 2 | 0 | 300 | 0 | False |
| 3 | ntl_sentinel_best_baseline | GDP per capita | 1 | 2 | 0 | 200 | 0 | False |
| 4 | sentinel_best_nowcast | GDP per capita | 2 | 2 | 0 | 300 | 0 | False |
| 5 | ntl_sentinel_best_nowcast | GDP per capita | 1 | 2 | 0 | 300 | 0 | False |
| 6 | sentinel_best_nowcast | GDP | 1 | 2 | 0 | 300 | -1 | True |
| 7 | ntl_sentinel_best_nowcast | GDP | 1 | 5 | 0 | 300 | -1 | True |

Notes: Rows 1-3 are related to the baseline model; Rows 4-5 pertain to the complete nowcasting model; Rows 6-7 represent the outcomes derived from the complete nowcasting model for GDP prediction. The train group contains data from 2018 to 2019.

## A1.3 Predicting Growth Rates

Table 1.15: Mean-squared error: GDP growth

| Row | Model | Label | Train MSE | Test MSE | Total MSE |
|---|---|---|---|---|---|
| 1 | Scenario 1 | GDPpc Growth | nan | 0.0680 | nan |
| 2 | Scenario 2 | GDPpc Growth | 0.0115 | 0.0297 | 0.0209 |
| 3 | Scenario 3 | GDPpc Growth | 0.0151 | 0.0241 | 0.0178 |
| 4 | Scenario 3 | GDP Growth | 0.0143 | 0.0222 | 0.0167 |

Notes: In Scenarios 1 and 2, the training group consists of data from 2019, while the test group comprises data from 2020. In Scenario 3, the training and test groups are split cross-sectionally, ensuring a balanced sample from all countries.

Table 1.16: Best parameters: three scenarios of predicting growth rates

| Row | Model | Label | ccp_alpha | max_depth | max_features | max_leaf_nodes | max_samples | min_impurity_decrease |
|---|---|---|---|---|---|---|---|---|
| 1 | Scenario 2 | GDPpc Growth | 0 | 0 | 37 | 0 | 0 | 0 |
| 2 | Scenario 3 | GDPpc Growth | 0 | 0 | 26 | 0 | 0 | 0 |
| 3 | Scenario 3 | GDP Growth | 0 | 0 | 13 | 0 | 0 | 0 |

Table 1.17: Best parameters: three scenarios of predicting growth rates (Continue)

| Row | Model | Label | min_samples_leaf | min_samples_split | min_weight_fraction_leaf | n_estimators | n_jobs | warm_start |
|---|---|---|---|---|---|---|---|---|
| 1 | Scenario 2 | GDPpc Growth | 4 | 2 | 0 | 200 | 0 | False |
| 2 | Scenario 3 | GDPpc Growth | 4 | 2 | 0 | 100 | 0 | False |
| 3 | Scenario 3 | GDP Growth | 4 | 2 | 0 | 100 | 0 | False |

Notes: In Scenarios 1 and 2, the training group consists of data from 2019, while the test group comprises data from 2020. In Scenario 3, the training and test groups are split cross-sectionally, ensuring a balanced sample from all countries.

# A2 Figures

## A2.1 Predicting GDP per Capita and GDP

Figure 1.16: Actual vs. predicted GDP, *ntl_sentinel_best_nowcast*

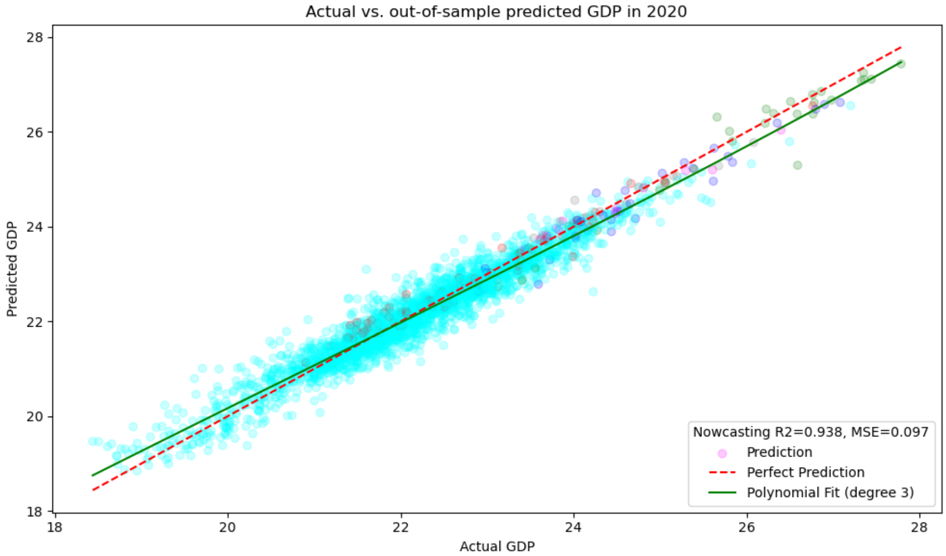



Figure 1.17: Important features, *ntl_best_nowcast*

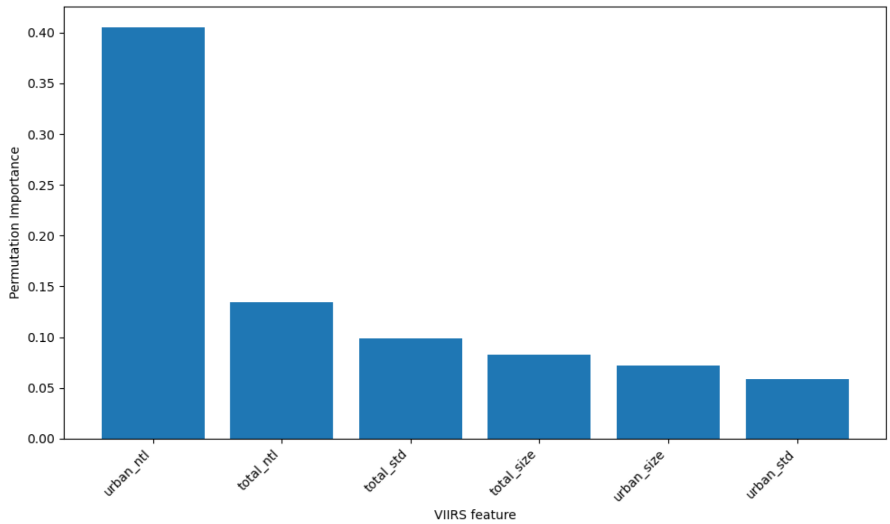


Figure 1.18: Important features, *sat_best_nowcast*

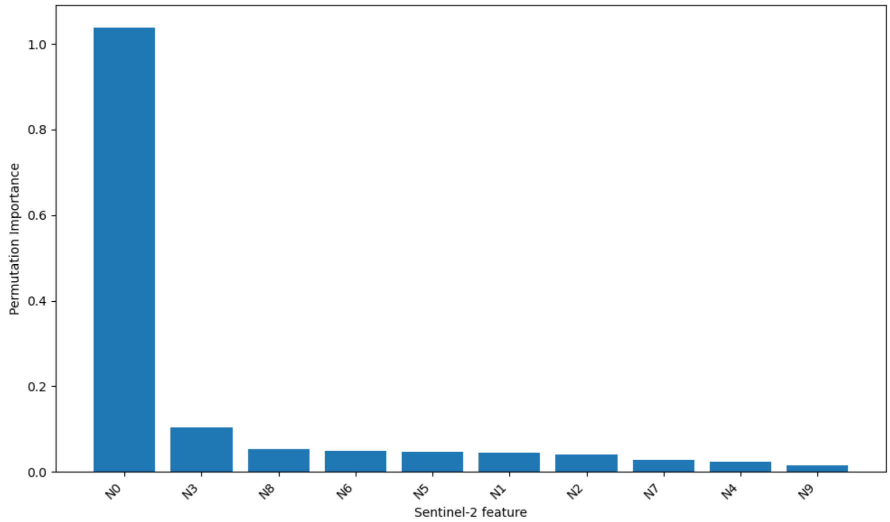
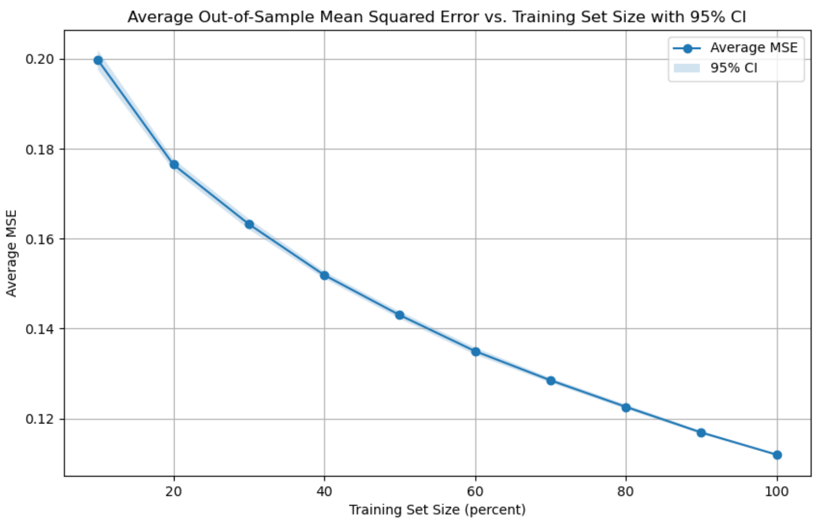


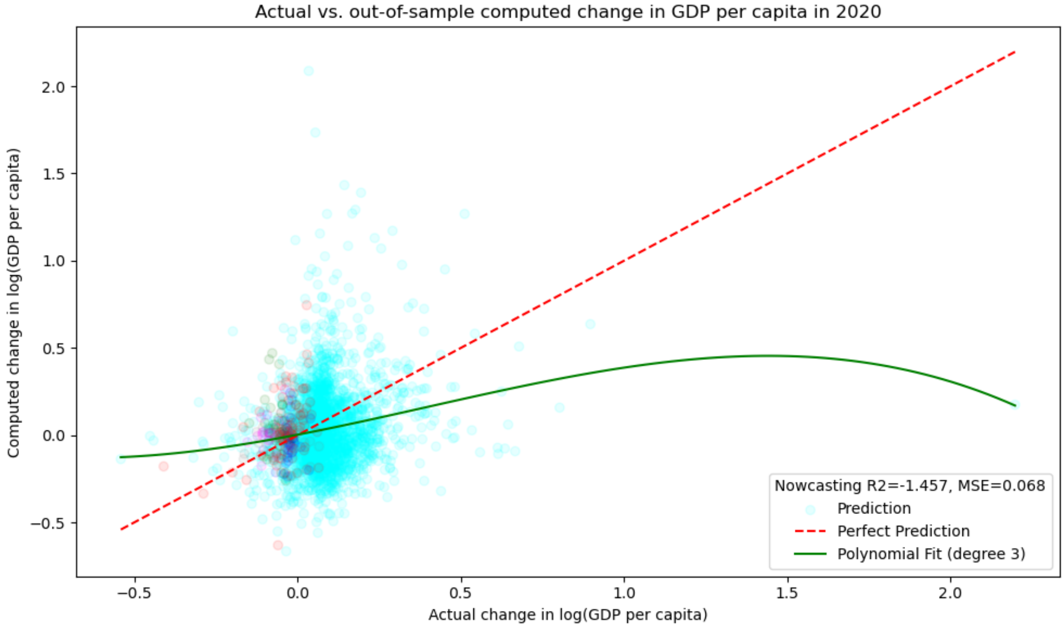

Figure 1.19: Mean squared error by size of training set, *ntl_sentinel_best_nowcast*




## A2.2 Predicting Growth Rates

Figure 1.20: Actual vs. computed change in GDP per capita

a) *Scenario 1*




b) *Scenario 2*

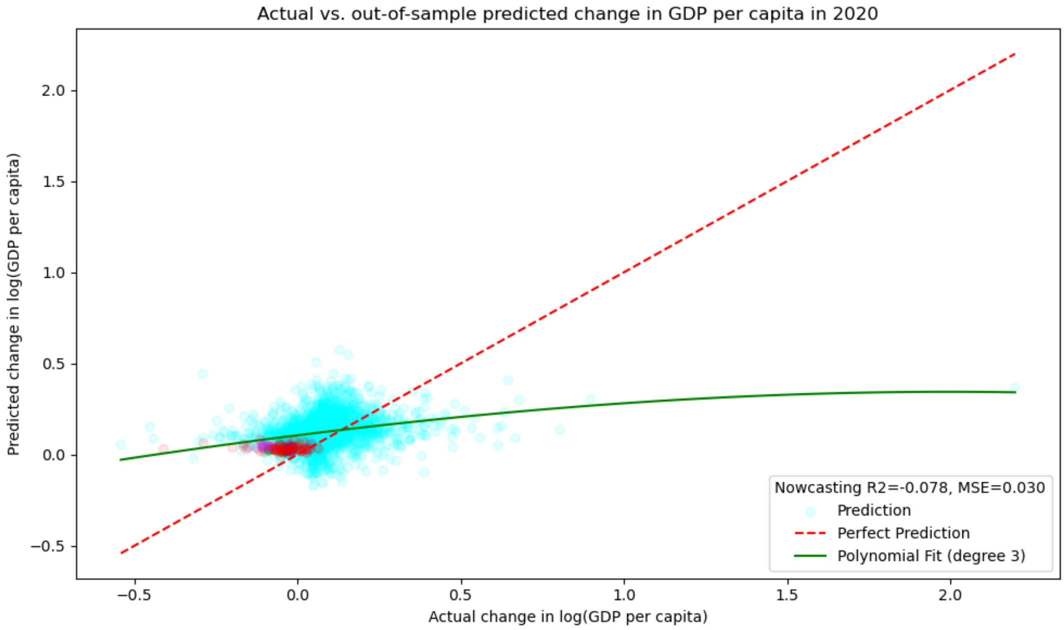
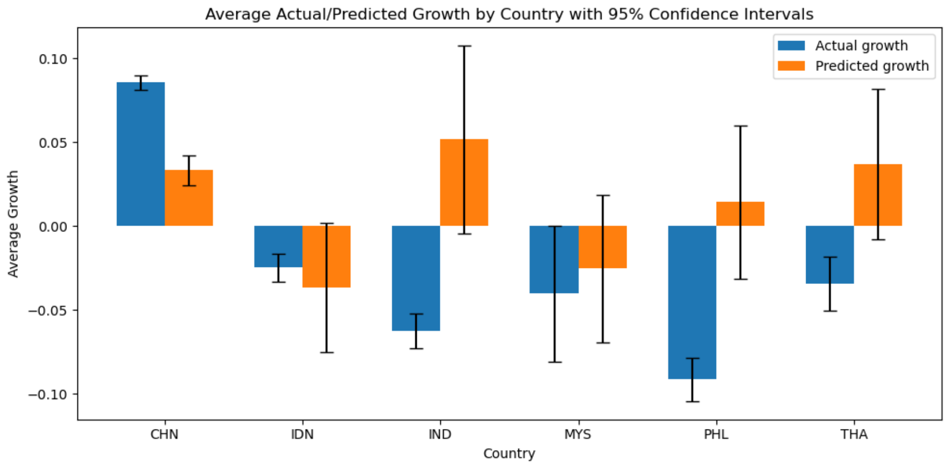


Figure 1.21: Average actual and computed growth rate by country with 95% confidence intervals

a) *Scenario 1*




b) *Scenario 2*

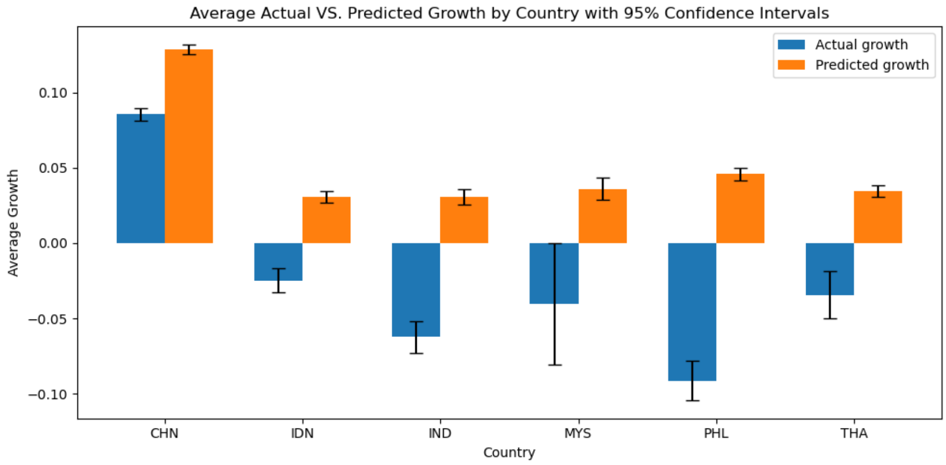


Figure 1.22: Actual vs. computed change in GDP per capita by country, *Scenario 1*

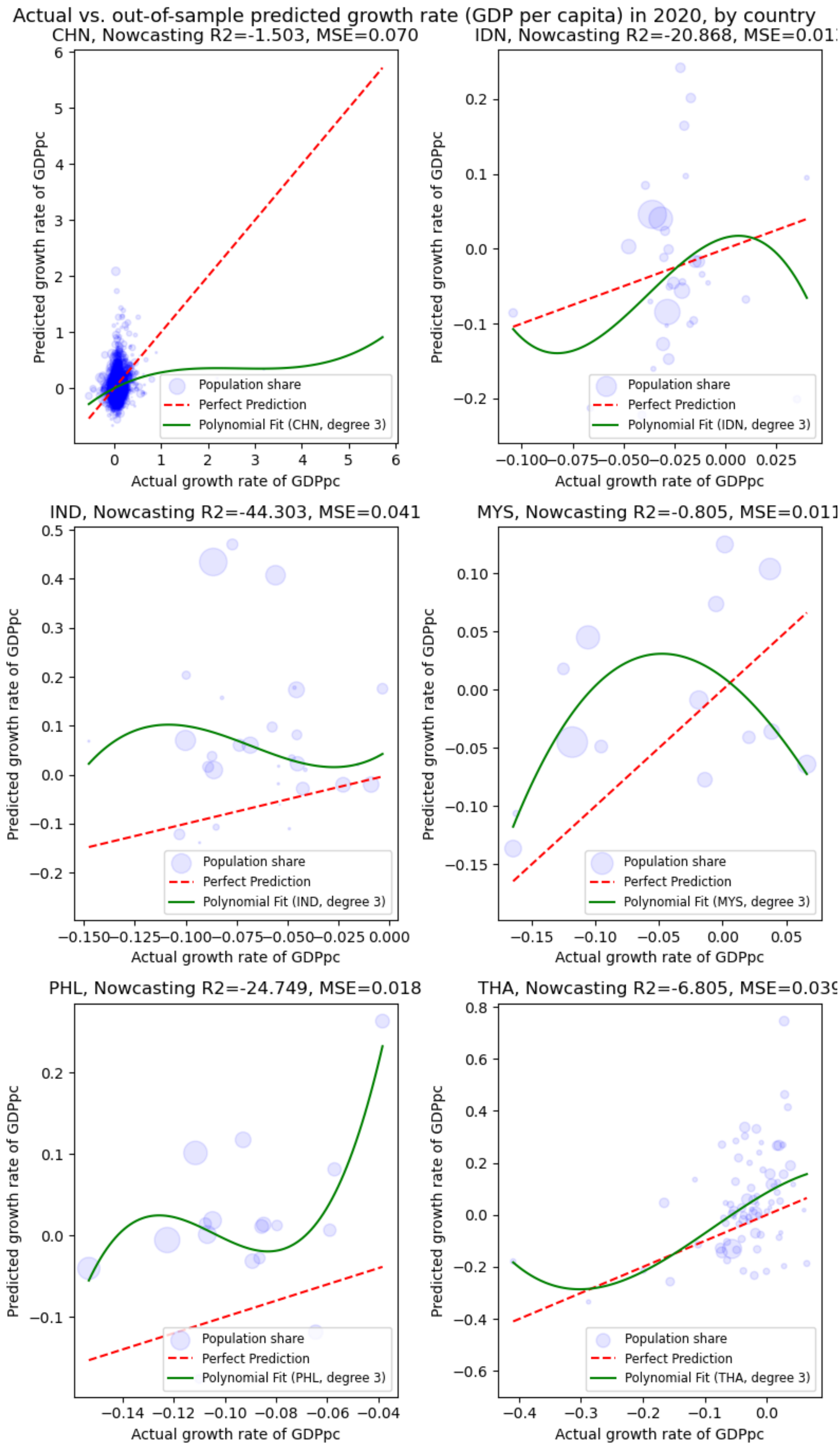


Figure 1.23: Actual vs. predicted change in GDP per capita by country, *Scenario 2*

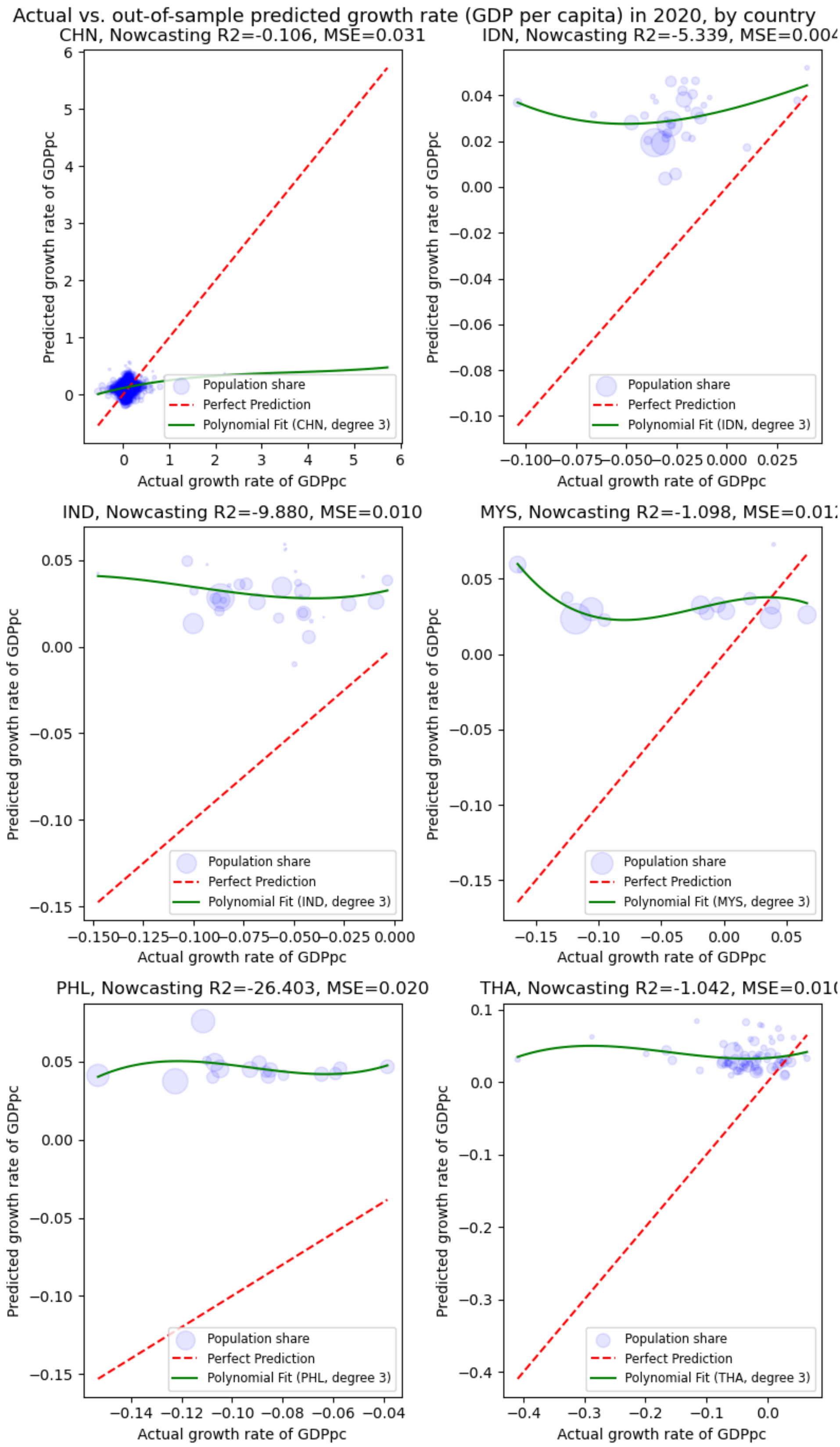
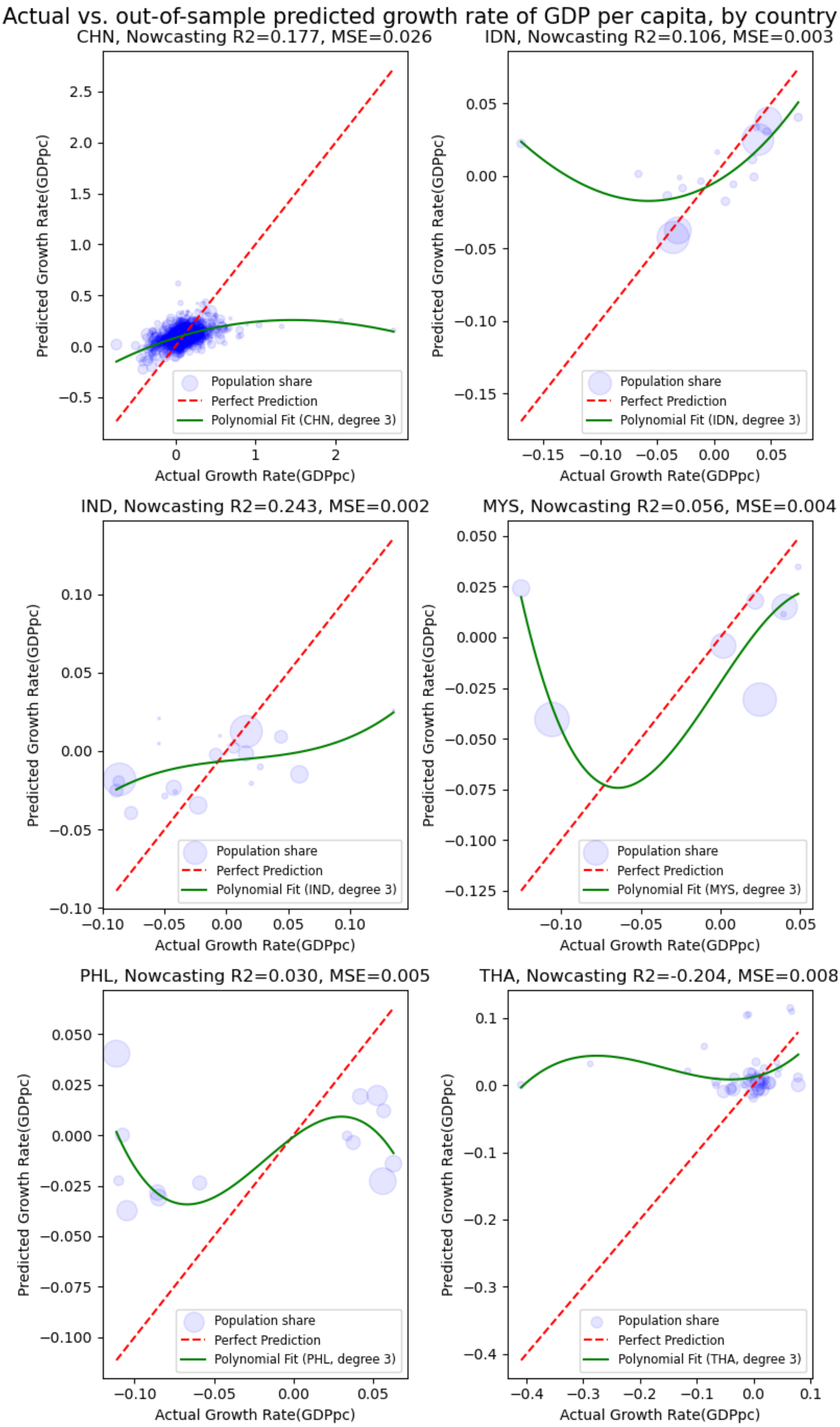


Figure 1.24: Actual vs. predicted change in GDP per capita by country, *Scenario 3*



# Chapter 2

# Conflict and Poverty[1]

## 2.1 Introduction

The goal of this chapter is to provide a detailed analysis of the link between poverty and armed conflict using the newest available conflict and poverty data both at the country level and at the regional level. We will first review the literature on the complex relationship between poverty, conflict, crime and other co-determinants. We then develop a measure based on the literature on the role of conflict risk and human capital accumulation, and use this measure to investigate the relationship between conflict and poverty.

Before discussing the role of armed conflict in relation to poverty, it is essential to acknowledge the limitations of this study. The analysis will assess the impact of conflict on poverty through two widely used datasets: the Uppsala Conflict Data Programme (UCDP) and the Armed Conflict Location & Event Data Project (ACLED), both of which track violent conflict globally. The UCDP employs a stringent definition of conflict, focusing on events resulting in battle-related deaths. According to this definition, battle-related deaths occur as a consequence of armed force between warring parties, whether state-based or non-state actors. This definition implies an emphasis on highly organized and extreme violence, excluding events such as riots. In contrast, ACLED adopts a more lenient definition of conflict events, while still providing coders with specific guidelines aimed at capturing political violence.

In our analysis, we relate conflict data to poverty at the country and regional levels, rather than examining household or neighborhood experiences or studying crime. This

[1]Coauthored with Hannes Mueller. We would like to thank Samuel Freije-Rodriguez, Michael Woolcock, Alexandru Cojocaru, Minh Cong Nguyen, Patrick Barron, and the participants of the Poverty, Fragility, and Conflict Webinar for their support, insightful comments, and suggestions.

constitutes a notable omission, as interpersonal violence, such as homicides, accounts for the majority of violent deaths. For instance, Hoeffler and Fearon (2015) highlight that for every civil war battlefield death, approximately nine individuals are killed in interpersonal disputes. The relevance of interpersonal violence and crime to economic activity is undeniable. However, as we will discuss in the following section, these effects do not manifest in macro-level economic changes but are better studied at the micro level, through data collection on directly affected individuals and firms. While we review the literature on these topics, our focus remains on the relationship between political violence and poverty.

In this chapter, we present a statistical framework designed to trace the cumulative long-term impacts of armed conflict on poverty, termed "conflict debt." Our findings reveal that contemporary conflicts create a conflict debt, which diminishes only gradually over time. The empirical model successfully captures both cross-country aggregate poverty patterns and regional cross-sectional variations. A key feature of the model is its acknowledgment that while armed conflict impedes poverty reduction, post-conflict periods may trigger substantial catch-up effects. Nevertheless, in countries most severely affected by conflict, recurrent cycles of violence continue to obstruct poverty recovery. Even in our most conservative estimates, the absence of conflict debt would result in poverty rates decreasing by 5-10 percentage points in these regions.

This article is structured as follows. We will first give a detailed overview over the existing literature on the economic costs of violence and explain the focus on armed conflict. We then discuss the cross-country evidence followed by regional evidence and a detailed case study of Uganda.

## 2.2 The Complex Relationship Between Conflict and Poverty

### 2.2.1 Violence and Poverty: An Overview

There is a clear pattern in the cross-country data - the poorest countries are also the countries which suffer most from armed conflict. This broad cross-sectional correlation could be driven by a lot of different factors including geography and colonial origins which affect both the tendency for conflict and poverty.

A first strategy to shed light on the causal link between conflict and poverty is to simply control for other factors at the macro level or to look at changes across time. The macroeconomic impact of civil wars has been studied extensively in this way for more than two decades and several studies have shown that countries which suffer from an armed conflict also suffer economic collapse. In one of the first studies, Collier (1999) analyses armed conflict and GDP growth using cross-county data on internal wars which occurred between 1960 and 1992. He finds that civil conflict is correlated with a contemporaneous reduction of GDP per capita growth of 2.2 percent. On a regional level, Abadie and Gardeazabal (2003) investigate the economic effect of the Basque terror campaign and estimate a reduction of the GDP per capita of approximately 10 percent. Some studies suggest even larger effects. Mueller and Tobias (2016) analyze economic growth around the onset of armed conflict and show that over a four-year period in conflict GDP per capita contracts by about 18% on average, relative to a counterfactual without conflict. This relative loss is not recovered once the conflict is over and more intense violence leads to larger economic declines. A detailed case study of the conflict in Syria by the World Bank conducted in 2017 finds that from 2011 until the end of 2016, the cumulative losses in gross domestic product (GDP) were USD 226 billion, about four times the Syrian GDP in 2010 (World Bank, 2017).

Economic recovery and poverty reductions might occur in the post-war period, but in many countries, peace does not last long enough. According to the definition of armed conflict we use in this report, there was at least some internal conflict in 27% of all country/years. But this violence has been particularly concentrated in a small set of countries. Within the last three decades, 12 countries have suffered more than three outbreaks of very intense armed conflict and another set of 8 countries suffered lasting, intense armed conflict in more than one third of the entire period. This means that many of the poorest countries never experienced long periods of stable peace. Observers have therefore talked of a conflict or violence trap which countries fail to escape.

Civil wars affect economic conditions in two stages: during conflict and during the post-war period. During conflict there is a direct cost caused by the destruction of resources that would otherwise have been employed in production. This directly impacts contemporaneous economic performance in the affected territory. In addition, the economic activity of countries that suffer from civil conflict is damaged indirectly by an increase in production costs and insecurity in transport. Risk perceptions play a key role here as they spread the costs of armed conflict beyond the directly affected territory. In addition, armed conflict will affect economic development and poverty also in its aftermath through the humanitarian crisis triggered by wars and the contraction of investments due to instability. A comprehensive review of all potential channels is beyond the scope of this chapter; however, four key issues are particularly pertinent and will be examined in detail:

- Armed conflict affects the productivity of the workforce in the long run through lack of education, stunting, injuries and mental disorders. It is as if conflict imposes a kind of *debt* burden that builds up during conflict and needs to be repaid once violent conflict has ended.

- Expectations, both temporal and spatial, play a crucial role. Fear of future outbreaks of violence will inhibit capital inflows and further reduce productivity. The fear of the spread of violence amplifies its effect beyond the directly affected individuals, firms and regions. Both of these factors increase the conflict debt.

- Crime affects economic outcomes. However, both in terms of human capital accumulation and the impact on expectations, political violence has far larger effects. This explains the much stronger link between poverty and political violence in the macro data.

- Weak state capacity will provide an additional link between a weak economy and internal conflict. Conflict is both a symptom of weak state capacity and a cause and so is a weak economy. This has repercussions for poverty alleviation strategies.

### 2.2.2 Evidence on the Human Capital Channel

Surprisingly, perhaps, there is quite a lot of controversy in the academic literature regarding the long-term effects of violence on economic activity. There are both theoretical and empirical reasons for this controversy. Theoretically, classic growth models like the Solow model assume that countries converge to their own steady state and that countries face decreasing returns to human and physical capital. An exogenous destruction of production factors through war will therefore result in a more rapid accumulation of physical or human capital after armed conflict. In its plainest form, this model predicts that countries will grow much quicker after a negative exogenous shock. The long-term economic costs of conflict are therefore less severe.

Empirically there is some support for this hypothesis from external wars. In an early contribution, Organski and Kugler (1977) show that output of winners and, more surprisingly, also losers of the World Wars caught up with non-affected countries after 15 to 20 years. This is supported by additional evidence from the micro level. Davis and Weinstein (2002) show that Japanese cities that were bombed had completely recovered their relative size twenty years after the US bombing in World War II. In a similar vein, Brakman et al. (2004) find that bombing of Germany had a significant but temporary impact on post-war city growth in West Germany but a sustained impact in East Germany. Minoiu and Shemyakina (2014) study a whole battery of indicators from poverty rates, population densities to schooling in Vietnam 25 years after the US bombing campaign and find that the local economy had recovered.

Of course, the time dimension involved in these findings is in itself not a reason for general optimism in a post-conflict situation. The macro evidence on quicker gains is more mixed - mostly because there is no agreed methodology for how to study recoveries. Cerra and Saxena (2008) use a method used to study crisis episodes more generally to study civil wars. In this method the start of an episode is coded as 1 and the impact over time is studied using forwards and lags. They find that, while output falls steeply immediately after conflict (6 percent on average), the economy recovers relatively soon afterwards. Mueller (2012) shows that this, fairly optimistic, view is due to a coding error. Economies contract by more than 18 percent in a civil war and this is not recovered within a decade. Mueller et al. (2017) confirm that there does not seem to be a distinct recovery growth period in the economic growth data. Countries appear to return to their average growth level in the absence of conflict. While this is good news in the sense that there is positive growth on average, it is also bad news in the sense that there seems to be, on average, no adjustment back to steady state once conflict is over. This is in contradiction with findings on the quick rebuilding of capital stock found in the literature. A possible theoretical explanation for this disparity between the micro and macro evidence comes from Barro and Sala-i Martin (2003) who predict that the speed of recovery depends on the type of capital that is destroyed, with a slower recovery if human capital is destroyed, because it has a higher adjustment cost.

This is particularly relevant as there is evidence in the literature that exposure to conflict during early childhood or adolescence has a huge and persistent direct effect on health, education and productivity of children. León (2012) shows that the average person exposed to political violence in Peru before school-age accumulated 0.31 fewer years of schooling upon reaching adulthood. This suggests that children who were affected by violence during early childhood suffer irreversible effects. Duflo (2001) finds that the effect of a huge school construction program in Indonesia on school attainment in the long-run is of a slightly smaller magnitude. Each school constructed per 1,000 children led to an increase of 0.12 to 0.19 years of education. This means that if one wanted to mitigate the long run effect of conflict in Peru on schooling by opening more schools, one would need to build more than 8,000 new schools. This is a big adjustment cost which supports the claims in Barro and Sala-i Martin (2003). Akresh et al. (2012) also identify the long-term impact of exposure to conflict during the Nigerian civil war in the late 60s on health. They find that individuals who were exposed to this conflict in their childhood or adolescence have a reduced stature. Adult stature is a latent stock measure of health and variation in height induced by childhood or adolescence environment has been documented in the literature as a good predictor of longevity, education and earnings (Behrman and Rosenzweig, 2004; Case and Paxson, 2008; Strauss and Thomas, 1998). The impact of pre-adulthood exposure to conflict on height identified in Akresh et al. (2012) can therefore be viewed as a permanent labor productivity loss for the exposed individuals.

The long-term effect working through human capital formation plays a key role for our study as it is consistent with the relationships between conflict and poverty in the country and regional data. This suggests that the latter are indeed driven by a causal relationship running from conflict to reduced economic capacity that is only slowly recovering. In this interpretation we can rely on a diverse body of evidence which finds large negative effects of armed conflict on education. Ichino and Winter-Ebmer (2004) for instance estimate the long run educational cost of World War II. They find that individuals who were 10 years old during the conflict in the most affected countries received less education than individuals with comparable characteristics in less affected countries. Affected individuals also experienced a sizable earnings loss, some 40 years later. Blattman and Annan (2010) exploit the exogeneity of child military conscription into rebel groups in Uganda to provide evidence of the extensive and long-lasting economic and educational impacts of conflict. Their findings indicate that schooling is reduced by nearly a year, skilled employment is halved, and earnings decrease by a third among former child soldiers compared to the control group. Akbulut-Yuksel (2014) exploits the plausibly exogenous region-by-cohort variation in the intensity of World War Two (WWII) destruction as a quasi-experiment and finds that exposure to destruction had long-lasting detrimental effects on the human capital formation, health, and labor market outcomes of Germans who were at school-age during WWII. An important channel for the effect of destruction on educational attainment is the destruction of schools whereas malnutrition is partly behind the estimated impact on health. Minoiu and Shemyakina (2014) use survey data and information on the dates and locations of conflict events to examine the causal impact of the 2002-2007 Ivorian conflict on child health. They find that children from conflict-affected regions suffered significant health setbacks and that victimization in the form of economic losses is the main conflict–impact mechanism.

Conflicts affect aggregate productivity directly through physical health but also indirectly through mental health. Being exposed to violence leads to post-traumatic stress disorder (PTSD) which affects both adults and children. These mental illnesses are characterized by hyper-vigilance, flashbacks and nightmares. Individuals who are exposed to a traumatic event have an increased likelihood of poor self-reported health, morbidity (as indicated by physical exam or laboratory tests), utilization of medical services, and mortality. Li et al. (2011) show that people who experience PTSD are at much higher risk of developing other health problems, including diabetes, heart ailments, depression and addiction. Sibai et al. (1989) study civilians exposed to the civil war in Lebanon and find that patients with severe coronary artery disease had more war-related stress than did either patients with normal arteriographic findings or hospital visitors. Using female Vietnam veterans, Wolfe et al. (1994) show that warzone exposure and PTSD are associated with self-reports of poor health and numerous physical problems. PTSD in early childhood also makes people more vulnerable in later life (Economist, 2015). Rosen and Fields (1988) were among the first to attempt to explain how PTSD could adversely affect physical health. The potential channels can be divided in 3 groups: biological, psychological and behavioral channels (Friedman and Schnurr, 1995). From a biological point of view, PTSD can modify cardiovascular reactivity, disturb sleep physiology, enhance thyroid function, etc. Singhal (2019) studies the long-term effects the American war in Vietnam on mental health. He finds that early-life exposure to war negatively affects mental health in adulthood and proposes malnutrition during the war as a significant mechanism.

We will see for poverty that poverty increases during conflict but that countries seem to catch up after conflict ends, i.e. that poverty rates converge towards a global average after civil wars. This rebound takes years even decades which is consistent with the idea that health and human capital are factors leading to poverty. In this view conflict leads to a loss in health and education in a generation of individuals and only once these individuals have partially recovered and re-built their human capital will they be able to escape poverty. The exposure of children to conflict is then what imposes the heaviest conflict debt we find in the data.

### 2.2.3 The Role of Expectations and Conflict Risk

The outbreak of a violence in a country will have a damning effect on expectations regarding renewed outbreaks of violence. The reason is that political conflict is, by and large, a re-occurring and spreading phenomenon. Violence spreads and always breaks out again in the same place. This effect is so powerful that research in conflict forecasting is struggling to develop variables that go beyond spatial and temporal dynamics when developing risk evaluations (Hegre et al., 2019; Bazzi et al., 2019; Mueller and Rauh, 2019). This is particularly relevant as the local population and outside economic actors also know this and when a new conflict breaks out it changes expectations dramatically.

Political violence is often unpredictable and in this way its impact on expectations is amplified. This is most clear for terrorism which actively tries to maximize its impact on expectations. Domestic violence and gang warfare might kill more people but unless extreme levels are reached the impact on expectations remains relatively low. Brazil and Mexico, for example, are regarded as great destinations for tourists despite extreme levels of violence which would put these countries firmly on the map of countries at civil war if the violence was politically motivated. Homicides are spatially contained, and the government, typically, has some control over their spread. Only when this is in doubt, i.e. when the violent actors turn against unpredictable targets or even the state, will economic actors react.

One indicator of how important fear as a factor is that the number of people running away from armed conflict is far larger than the number of fatalities. Mueller et al. (2017) show that in the average civil war year over 600,000 individuals have fled the country. Many more are displaced within country. Fear will therefore break local supply chains. Amodio et al. (2020) show, for example, that conflict during the Second Intifada in Palestine led to distortions in the functioning and accessibility of markets for production inputs, inducing firms to substitute domestically produced materials for imported ones. These distortions explain over 70 percent of the drop in the output value of firms in high-conflict districts. Based on these observations, Tapsoba (2018) proposes metric that captures perceived violence risk on the ground and uses this metric to evaluate the impact of violence risk on child health. His proposed measure of violence risk is based on an estimation of the underlying distribution of the violence process in space and time. The empirical results show that cohorts of children exposed to high risk of violence suffer major health setbacks even if they were not directly exposed to violence. In particular, risk increases infant mortality by more than 50% in both the Ivory Coast and Uganda. This means that the effects of armed violence are amplified far beyond what can be measured in the regions which are immediately affected by violence.

An important mechanism for the spread of fear and amplification of the effect of violence is news coverage. Besley et al. (2020) show at the example of tourists traveling to destination countries suffering from violence that a large part of the effect of violence is coming from news reporting on this violence in the home countries of tourists. Tourists from countries that report a lot on violence are much more likely to stay away. If this mechanism generalizes then reporting on countries in armed conflict will generate a lot of bad press and this will lower business trips and cut countries or regions off from trade and investment, i.e. slow economic integration in the long run.

But there is also a temporal aspect to fear. Asset prices can be used to show how expectations change with political violence. Zussman and Zussman (2006) and Willard et al. (1996) show, for example, that asset prices during conflict react to important conflict events like battles or cease fire agreements. Besley and Mueller (2012) show that house prices seem to react to changes in expectations rather than violence itself. This implies that even once violence is over it can be that asset prices stay suppressed. Parts of the economy will then only recover when peace is regarded as stable. This generates huge problems in fragile states which, as the word suggests, are not regarded as stable. Poverty reduction in these countries might rely as much on changing perceptions of political stability as on economic help.

This fear effect also has a strong impact of how the effect of local violence should be captured at the macro level. If the effect of violence is only contemporaneous then we do not expect conflict to affect poverty beyond the conflict period. However, if expectations are important, we expect economic effects to persist as with a debt that needs to be repaid after conflict has ended. Expectations will then also lead to the sort of patterns we find in the poverty data.

### 2.2.4 Poverty and the Cost of Crime

There is no robust empirical relationship between homicide rates on growth rates at the macro level, neither for regions within counties nor across countries. This is puzzling given the extreme levels of violence associated with crime in some regions. A possible reason for this lack of a relationship is that crime does not affect expectations in the same way as political violence. Also, crime rates move much more slowly and are therefore not causing the same dramatic disruptions or refugee waves as political violence.

This does not mean that violent crimes might not have a profound impact on the economy. There is a large related literature that calculates the cost of crime (See Soares (2009) for a review). The standard accounting approach is to estimate this cost simply by adding the costs and losses due to crime. For example, van Ours and Vollaard (2016) use estimates from an accounting approach to study the welfare gains from installing electronic engine immobilizers in the Netherlands. Their benchmark measures of the cost come from the UK Home Office. The methodology builds on Brand and Price (2000), a detailed accounting exercise that aggregates security expenditures, insurance costs, and damages to derive a per-case cost. Alternative approaches include individual valuations of counterfactuals, contingent-valuation, and changes in market prices to estimate welfare costs. For instance, Cook and MacDonald (2011) employ both contingent-valuation surveys and jury awards to victims of violent crimes to estimate the welfare gains from crime reductions. Similarly, property prices are used by Gibbons (2004) to assess the cost of crime in London.

Recent research on the economic effects of crime have shown that many of these studies will still constitute a lower bound of the cost of crime. Besley and Mueller (2018) show, for example, that two thirds of the economic impact of crime on firms comes through defensive measures taken by firms which might or might not be directly affected by crime. In other words, if one would focus on crime statistics to understand the impact of crime one would only capture one third of its actual costs. Crime might disappear where firms are successful defending against it. But the effort exerted in defending against crime forms part of the cost.

There is therefore little doubt that crime will affect poverty rates through its impact on the allocation of jobs across firms and the arising disincentives for investment that arise from the lack of secure property rights. But this effect will only be visible in individual and firm level surveys or by calculating costs of criminal acts directly.

### 2.2.5 Fragility, State Capacity and Poverty Reduction

Fragility will affect expectations and suppress investment incentives and, in this way, poverty. But the mechanism running between conflict and poverty are much more complex. This becomes clear when one looks at the six symptoms proposed by the International Growth Centre (IGC) in a recent study on fragility (The LSE-Oxford Commission on State Fragility and Development, 2018):

1. Weak state capacity: a failure to invest in fiscal, legal, regulatory and spending capabilities of government;

2. A weak private sector: economies are characterized by a large informal sector with few large firms and poor legal structures in place, which hampers taxation;

3. Lack of security: the state is not able to provide security from disruptive actors (such as organized criminals or militias) throughout its territory;

4. Weak resilience: the economy often relies on few sectors and is subject to external shocks which threaten political stability;

5. Low levels of state legitimacy: society suffers from low levels of trust and reciprocal compliance; and

6. Polarized societies are characterized by the prevalence of oppositional identities, whether ideological, ethnic, linguistic, or religious. Different polities exhibit these traits to varying degrees. Moreover, countries that have functioned effectively for extended periods may intermittently display some of these symptoms.

From this it is clear that fragility and poverty are also linked through weak state capacity. The inability of the government to provide public services will, for example, directly affect its ability to combat poverty. This implies that the causal chain might run from conflict to human capital formation as proposed above but it might also run from low state capacity to conflict and low human capital formation. Empirically, it is therefore difficult to pin down a specific channel. We therefore rely on the micro literature discussed above to establish the causal link.

But this view on fragility should be taken into account when designing policy for poverty alleviation. Clearly, poverty reduction must partly come from reducing fragility and regaining state capacity. Besley and Mueller (2020) argue that crafting a program for building fiscal capacity to suit the context of state fragility means tailoring the narrative of reform to the social and cultural context. For example, poverty alleviation in fragile states must recognize the lack of legitimacy stemming from poor accountability in the use of public resources and corruption by elites. This needs to become a standard part of understanding the context for building state capacity for sustained poverty reductions. A country with low trust in government, weak national identity, and low confidence in the state may need a different approach from that taken where these factors are not present.

An important case study in this context comes from Sanchez de la Sierra (2020). He shows that stationary armed non-state actors in the Democratic Republic of Congo seem to improve economic welfare for local villages whereas the Congolese army only increases security. The reason is that other public services by the central state are not provided in remote areas. In such a context one would expect poverty measures to remain high with the presence of the army. This, in turn, will weaken support for the state and state revenues because the formation of norms of reciprocity rely on citizens seeing that the state is using revenues for common purposes. The presence of strong, local ethnic or regional identities can make this even more important.

An approach to poverty alleviation in fragile states that recognizes the need to strengthen the social contract therefore reinforces the need for organizations such as the UN, World Bank and IMF to coordinate when strengthening state capacity. Programs may work better if there is scope to coordinate efforts to improve security (which means coordinating with security operations) with projects to improve service delivery.[2]

## 2.3 Data

### 2.3.1 Poverty Data

Throughout this report, we utilize three poverty datasets to examine the relationship between poverty and conflict. The first dataset is an unbalanced panel of the poverty headcount ratio at $1.90 per day (2011 PPP) at the country level, sourced from the World Bank's PovcalNet. The second dataset is the Global Subnational Atlas of Poverty (GSAP), which provides a cross-sectional overview of international poverty measures around 2013 at the subnational level. The third dataset is a short panel from Uganda, covering the years 2012 and 2016, comprising approximately 1,500 sub-counties in the country.

[2]The case of Egypt illustrates the complexities of implementing reforms in practice. Fiscal reforms were accompanied by a World Bank project aimed at improving social security. However, the shift in government following the removal of Morsi undermined the project, as the new administration sought to consolidate control. National data indicate that poverty has since increased. Ironically, sharing fiscal resources with the population can also serve as a means of reinforcing fiscal control.

### 2.3.2 Conflict Data

Two conflict databases are used in the study. The Georeferenced Event Dataset (GED) Global version 19.1 from the Uppsala Conflict Data Programme (UCDP), henceforth called UCDP, and the Armed Conflict Location & Event Data Project (ACLED) are databases of geolocated conflict events around the world with a number of fatalities. The two datasets are different in terms of coverage and data collecting process. The UCDP has larger coverage with more countries so it delivers a broader illustration of the global situation between 1989-2018 when we implement the cross-country study. But UCDP's different inclusion criteria, such that a conflict dyad is only coded once it crosses the 25 battle-related deaths threshold in a given year, makes it understate fatalities from conflict in many areas (see Raleigh and Kishi 2019 comparing the two datasets in details). When we analyze the within-country relationship between poverty and conflict, ACLED is providing a very rich set of data and we therefore show results for our within-country case study and the African continent using the ACLED data.

Table 2.1: Summary Statistics, Conflict Debt and Poverty Rate

| | N | Mean | SD | Min | Max |
|---|---|---|---|---|---|
| Countries | | | | | |
| Poverty rate (%) | 3,238 | 17.26 | 22.80 | 0.00 | 94.10 |
| Conflict debt | 3,238 | 0.18 | 0.61 | 0.00 | 4.97 |
| Subregions | | | | | |
| Poverty rate (%) | 1,885 | 15.98 | 24.49 | 0.00 | 97.55 |
| Conflict debt | 1,885 | 0.09 | 0.36 | 0.00 | 3.98 |
| Subregions in Africa | | | | | |
| Poverty rate (%) | 618 | 40.20 | 27.26 | 0.00 | 97.55 |
| Conflict debt | 618 | 0.14 | 0.43 | 0.00 | 3.11 |
| Conflict debt (ACLED) | 618 | 0.18 | 0.54 | 0.00 | 3.57 |
| Sub-counties in Uganda | | | | | |
| Poverty rate (%) | 3,026 | 22.65 | 16.25 | 0.61 | 94.16 |
| Conflict debt | 3,026 | 0.07 | 0.25 | 0.00 | 3.09 |

Notes: All violence data are from UCDP, except subregions in Africa when the data from ACLED are used for comparison. The poverty rate is the poverty headcount ratio at $1.90 a day (2011 PPP). Countries group is a cross-country panel during 1989-2018. The poverty data of this group are interpolated linearly to replace missing values between two data points. The group of Subregions is cross-sectional from the Global Poverty Geodatabase 1992-2018. Subregions in Africa are from the previous data 2008-2018. The sample of Sub-counties in Uganda is from 2012 and 2016.

## 2.4 The Empirical Relationship between Conflict and Poverty

In what follows we will take an agnostic view regarding the channel through which the correlation of violence on poverty rates arises. Instead, we will rely on the related literature for identification and simply try to understand the involved magnitudes in macro relationship at the country and regional level. If conflict prevents children to build their human capital or hinders investments in the recovery period, then past conflict should be related to poverty rates today. The channels of human capital accumulation and expectations will then lead to a correlation between poverty rates and past conflict. Theoretically, we expect reverse causality from poverty to conflict to generate a contemporaneous relationship. If poverty triggers conflict, then increases in poverty should increase contemporaneous conflict risk.

We will assume that the relationship between violence and poverty works through the local level and re-code conflict counts into per capita counts, i.e. divide fatalities by the overall population of the country or region. This assumes that the effect of violence on the poverty rate is more severe if a conflict is affecting a larger share of the population. In a country like India a fatality count will mean that a much smaller share of the population is affected than in a small country like Burkina Faso. Fatality counts have to be seen in the context of the total population. This also true for regions. Violence in a region with a large population will have smaller effects on the economy per capita. An important feature of this way of coding violence intensity is that we can compare this intensity across units of very different sizes.

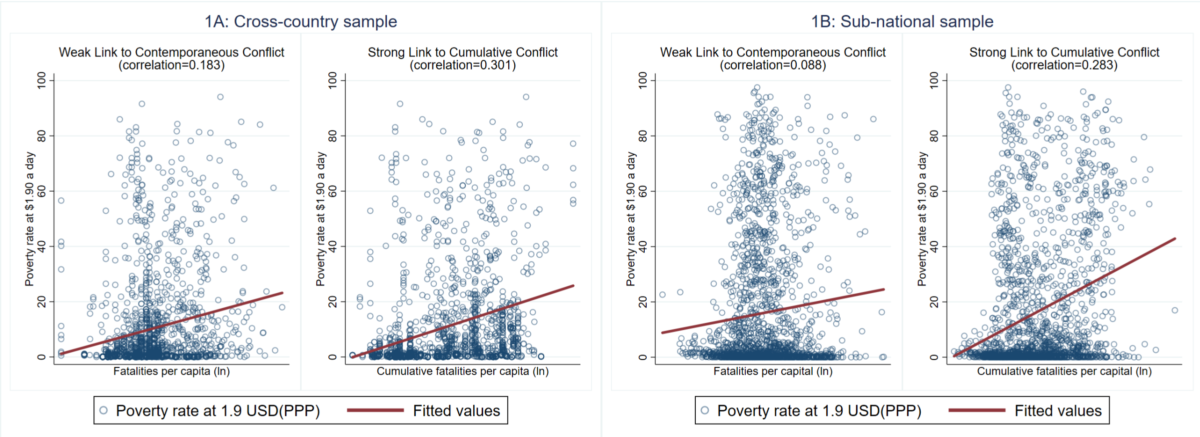
The per capita measure is, in essence, a theoretical statement on the likely link between conflict and poverty. It is, for example, consistent with the literature discussed in the previous sections which shows that conflict inhibits human capital formation locally. The higher the share of children affected by conflict the lower will productivity be in future.

### 2.4.1 The Conflict Debt Measure

Figure 2.1 shows simple correlations between per capita fatalities and the poverty rate using the cross-country and sub-national samples in Figure 1A and 1B, respectively. In the Figures we distinguish between contemporaneous conflict which is the (log of the) fatality count in the same year as the poverty measure and the cumulative count which is the (log of the) total fatality count between 1989 and the year of the poverty measurement. The idea behind producing the total past count is that past conflict could affect poverty today through the channel of human capital accumulation, for example. If violence a decade ago harmed children growing up at the time, then this could have an impact on the ability of those individuals to gain a living ten years later.

Clearly, we see that there is a positive relationship between both contemporaneous and past conflict and contemporaneous poverty rate in both panels of Figure 2.1. More intense conflict, present and past, is strongly associated with higher poverty rates at present. However, we also see that the link between contemporaneous conflict and poverty is a little weaker.

Figure 2.1: Relationship between contemporaneous and past cumulative conflict and poverty



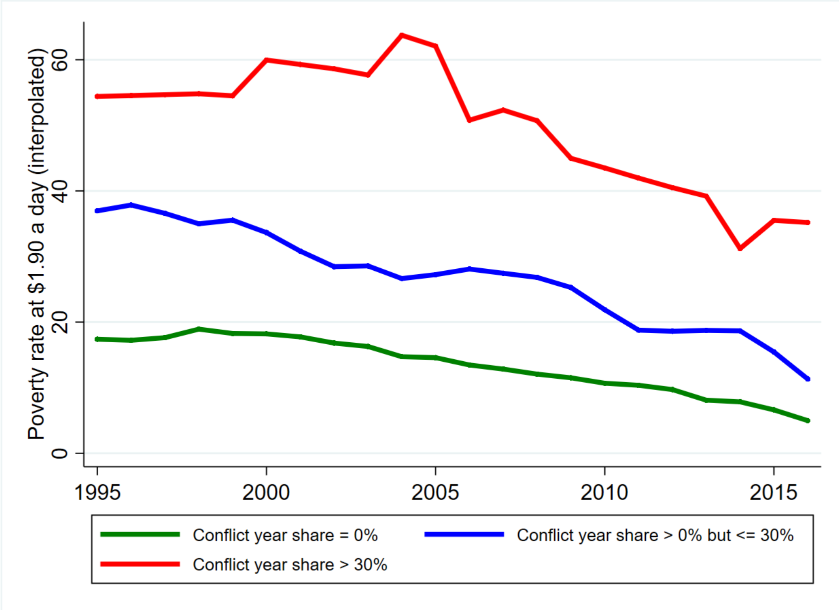

Sources: PovcalNet and GSAP by World Bank, and UCDP

In what follows we adopt the definition of conflict year as having fatalities per 1,000 population more than 0.08 to define a country as being in conflict or not. Mueller (2016) proposes this threshold for conflict intensity as this is about as common as civil war defined as a year with more than 1000 fatalities (the definition typically used in the conflict literature). We experimented with the threshold and with absolute counts and the latter is indeed only weakly related to the poverty rate. Cumulative conflict years as defined here are strongly related to poverty. This is in line with the idea that conflict prevents countries from eradicating poverty as public resources are scarce, capital takes flight and as the population suffers human capital losses. These losses scar the economy of the country and past conflict therefore affects poverty today.

However, there is a dynamic aspect to poverty reduction which a simple cumulative measure does not capture. Note, that a cumulative conflict measure implies that a year of conflict ten years ago will have the same impact on the economy as a year of conflict five years ago. If the impact weakens over time, for example because individuals recover from conflict, then this way of capturing conflict history will be an empirical misspecification. We turn towards this in Figure 2.2.

Figure 2.2 shows the average poverty rate in all countries with poverty data for the period 1995 to 2016 but divides countries into three groups according to their conflict

Figure 2.2: Poverty rates across time




Sources: PovcalNet by World Bank, and UCDP. Notes: A conflict year is defined by a year with over 0.08 fatalities per 1,000 population. We use average population to avoid time variation in the denominator of the indicator. Average population is computed from available years in the World Bank data. The conflict year share is a share of conflict years in the period 1990-2019.

history.[3] The first group of countries (green line) is not affected by conflict at all. This group includes most developed countries. The second group experienced conflict but in less than 30% of all years (blue line) and the third group of countries experienced conflict in more than 30% of all years (red line). This last group contains countries like Afghanistan and Somalia which are heavily affected by repeated cycles of armed conflict.

If we focus on the most affected countries and compare to the countries which are not affected, we see a clear drifting apart until around 2005 and then a slight catching up. Poverty rates in the countries most affected by conflict remain extremely high with the overall average still close to 40 percent in 2016. This implies that the countries which were most heavily affected by conflict did not manage to combat poverty effectively. Clearly, what we can see here is the "Two-Speed World" invoked by Corral et al. (2020) where some conflict countries are actually converging with the rest of the world in terms of poverty numbers and the most affected countries are left behind. If we contrast countries which had some conflict (blue line) and no-conflict countries (green line), it is, however, not clear whether some conflict inhibits poverty reduction. In fact, from this comparison it seems as if the conflict countries might be catching up with the non-conflict countries. This is not surprising as the conflict countries include countries like Colombia, Uganda

[3]The poverty data is interpolated to generate data which is as consistent as possible across time. We rely on our definition as a conflict year as being a year as having more than 0.08 fatalities per 1,000 population and the sample are 187 countries in the period 1989 to 2018 to define groups.

and Angola which all are increasingly peaceful.

This observation has an extremely important implication: conditional on the absence of contemporaneous or recent conflict there seems to be some convergence in poverty rates. This finding is further supported by Corral et al. (2020), who demonstrate significant poverty reduction in regions that have escaped "fragile and conflict-affected situations." We therefore need to use a more nuanced framework to understand the dynamic effects of conflict on poverty reduction. A simple cumulative conflict history without discounting will not be a good representation of the fact that countries seem to be able to escape high poverty rates once conflict ends.

To empirically capture this relationship, we construct a model of conflict debt and its impact on poverty rates, represented by the following two equations:

$$PovertyRate_{it} = \beta\ ConflictDebt_{it} + \mu_i + \mu_t + \varepsilon_{it}, \tag{2.1}$$

where $\mu_i,\ \mu_t$ represent country and time fixed effects, respectively.

$$ConflictDebt_{it} = \delta\ ConflictDebt_{it-1} + ConflictYear_{it}, \tag{2.2}$$

where $ConflictYear_{it}$ is a dummy variable equal to 1 in years with more than 0.08 fatalities per 1,000 population.

Equation 2.2 captures the build-up of a stock of conflict debt with ongoing conflict. $ConflictDebt_{it}$ increases by 1 in a conflict year and increases by $\delta$ with past conflict debt. The parameter $\delta$ is a decay parameter which will be smaller than 1 if conflict in the past affects current poverty less and less over time, i.e. if countries are able to catch up and "repay their debt" after conflict is over. In the context of the literature discussion the parameter $\delta$ is best thought of as the (mental) health and skills of the affected population recovering in the years following conflict or as investors regarding a region or country as increasingly stable with lasting peace.

We follow a two-stage procedure. We first calculate the value of $ConflictDebt_{it}$ for every country year under different assumptions on $\delta$. We then estimate $\delta$ by running the regression in (1) repeatedly with different values of $\delta$ and taking the regression that maximizes the fit of the model. As a result, an estimated $\delta$ equal to 0.88 is used throughout the study. This implies that, with stable peace, a conflict debt shrinks to 52% or its original size after 5 years and to 28% of its original size in 10 years. This is in itself interesting as it suggests a relatively quick recovery of poverty rates after conflict ends.

In Table 2.1 we show the evaluation of the conflict data using the conflict debt model. In the cross-country data we show that the mean debt value is 0.18 but debt concentrates in few countries. Conflict debt takes a maximum value of almost 5 as the debt accumulates without peace.

### 2.4.2 Cross Country Results

The results provided by our conflict debt model are shown in Table 2.2 below. There is a robust relationship between conflict debt and the poverty rate where an increase by 1 in the debt is associated with an increase in the poverty rate of 1.77 percentage points in our preferred specification in Column (3) which features country and year fixed effects. What is important to note here is how the country fixed effects reduce the size of the coefficient. This suggests that level differences in conflict and poverty across countries are a key driver of the relationship observed in Figure 2.1A. Once these differences are accounted for, the strength of the relationship diminishes significantly.

The statistical robustness to controlling for country fixed effects is an important reassurance that the relationship between conflict history and poverty is not solely due to a host of omitted variables at the country level. Controlling for continent/time fixed effects does not alter the size of the coefficient substantially. This implies that poverty in some countries deviates from the regional average in line with conflict debt. Still, we will need to interpret the coefficient carefully, keeping in mind that the link between conflict and poverty could be driven by other factors which affect both.

Table 2.2: Poverty Rate and Conflict Debt

| | (1) Poverty rate | (2) Poverty rate | (3) Poverty rate | (4) Poverty rate |
|---|---|---|---|---|
| Conflict debt | 8.372*** | 7.971*** | 1.767** | 1.455* |
| | (0.970) | (0.943) | (0.877) | (0.856) |
| Country fixed effects | No | No | Yes | Yes |
| Time fixed effects | No | Yes | Yes | No |
| Continent-time fixed effects | No | No | No | Yes |
| Adjusted R-squared | 0.0502 | 0.0800 | 0.277 | 0.421 |
| Observations | 3238 | 3238 | 3238 | 3238 |

Notes: Robust standard errors clustered at the country level in parentheses. * $p<0.1$, ** $p<0.05$, *** $p<0.01$. The dependent variable is poverty rate (in percent), interpolated linearly to replace missing values between two data points. Conflict debt is the discounted history of conflict years where a conflict year is defined as by more than 0.08 fatalities per 1,000 population. We use average population to avoid time variation in the denominator of the indicator. Average population is computed from available years in the World Bank data. Continent-time fixed effects are a dummy for every year on every continent. Continent here are 7 world regions. Regressions are on unbalanced cross-country sample during 1989-2018.

An additional, immediate concern in Table 2.2 is reverse causality where poverty leads to conflict and not the other way around. And indeed, the robust relationship between conflict debt and poverty rates shown in Table 2.2 is still present if we use a simple dummy for contemporaneous conflict as shown in Table 2.3 . However, in

Appendix Table 2.7 we show that the association between conflict debt and the poverty rate is robust to controlling for contemporaneous conflict whereas the reverse is not true. Countries with large conflict debts suffer from poverty but, controlling for this debt, there is no association between poverty and armed conflict. Also, in Column (4) of Table 2.3 , we see that contemporaneous conflict loses its size and significance when controlling for continent time fixed effects. This means that poverty rates are stronger related to past conflict than to present conflict which implies that reverse causality cannot be the main driver of the correlations shown in Table 2.2.

Table 2.3: Poverty Rate and Conflict

| | (1)<br>Poverty rate | (2)<br>Poverty rate | (3)<br>Poverty rate | (4)<br>Poverty rate |
|---|---|---|---|---|
| Conflict debt | 8.372*** | 7.971*** | 1.767** | 1.455* |
| | (0.970) | (0.943) | (0.877) | (0.856) |
| Country fixed effects | No | No | Yes | Yes |
| Time fixed effects | No | Yes | Yes | No |
| Continent-time fixed effects | No | No | No | Yes |
| Adjusted R-squared | 0.0502 | 0.0800 | 0.277 | 0.421 |
| Observations | 3238 | 3238 | 3238 | 3238 |

Notes: Robust standard errors clustered at the country level in parentheses. * p<0.1, ** p<0.05, *** p<0.01. The dependent variable is poverty rate (in percent), interpolated linearly to replace missing values between two data points. A conflict year is defined by a year with 0.08 fatalities per 1,000 population. We use average population to avoid time variation in the denominator of the indicator. Average population is computed from available years in the World Bank data. Continent-time fixed effects are a dummy for every year on every continent. Regressions are on unbalanced cross-country sample during 1989-2018.

There is a significant intensive margin in the relationship between conflict debt and poverty. A heavier debt burden is correlated with higher poverty rates, even when focusing solely on country-years with conflict debt. The greater the debt burden, the higher the poverty levels, as shown in Appendix Table 2.8. This also indicates the presence of strong catch-up effects as the debt diminishes, which will be further discussed in the following sections.

The second aspect is that the link we find is not specific to the poverty measure we use. In the appendix we use the multi-dimensional poverty index developed in Alkire and Jahan (2018) to show that the strong link between debt burden and poverty also holds in that data (Appendix Table 2.9). We find extremely similar magnitudes to the results shown in Table 2.2. The multidimensional aspect of the data also allows to investigate the channels through which conflict debt might affect conflict. We find very little on the health dimension (nutrition and child mortality) but strong links to the education dimension and even stronger links to all standard of living categories. This is reasonable as it is less likely that, years after conflict ended, conflict history still affects harvests. But if it disrupts the economy and affects human capital accumulation then this will affect years of schooling and assets in the mid-term. We find that deprivation of public services such as sanitation, electricity, and drinking water increases with the conflict burden, suggesting that state capacity may be a relevant confounding factor.

When interpreting Tables 2.2 and 2.3, it is important to recognize that the aim of our empirical analysis is not to demonstrate the dynamic response of poverty to innovations in conflict. Rather, our focus is on quantifying broad trends at the country level, as illustrated in Figures 2.1 and 2.2. Our empirical results indicate that poverty trends align with a model in which conflict generates a persistent debt. We will apply this model to cross-sectional data below, where we observe the strength of our statistical approach. It offers valuable hypotheses for both panel and cross-sectional data at various levels of analysis.

If we interpret the coefficient in Column (3) of Table 2.2 as causal, the economic significance of the relationship can be evaluated as follows: The standard deviation of the conflict debt variable is 0.6, which suggests that an increase of one standard deviation in conflict debt is associated with an approximate 1 percentage point increase in the poverty rate (0.6 * 1.767). However, using the standard deviation as a measure of the burden may underestimate the effect in some countries, as the maximum value of the conflict debt variable is 5. This would correspond to an increase of nearly 9 percentage points in the poverty rate.

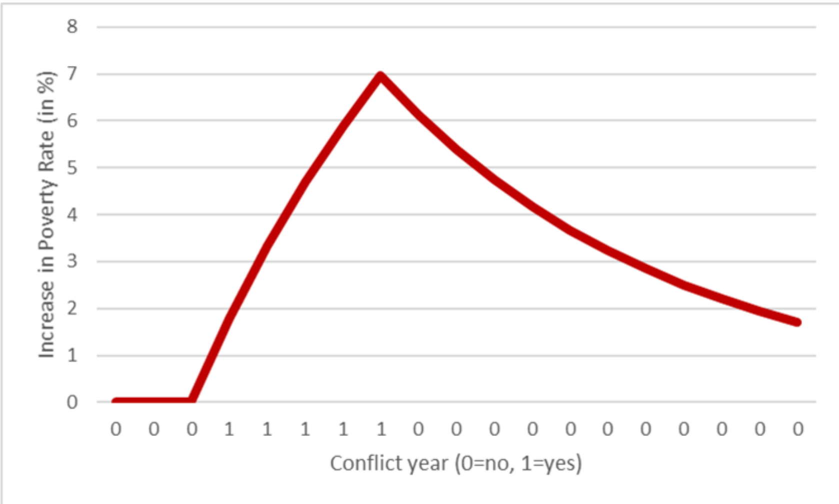
One way to understand the dynamic relationship identified in the regression is to simulate conflict debt for a conflict and plot the implied divergence in poverty rates over time. In Figure 2.3, we illustrate the projected increase in the poverty rate during a conflict lasting five years, represented by a series of five 1s on the X-axis. As the conflict begins, the poverty rate rises sharply, reaching over 7 percentage points higher by the end of the five-year period. Although the poverty rate gradually decreases as conflict debt diminishes, it remains approximately 1 percentage point higher even ten years after the conflict ends.

This pattern suggests a form of "catching up," as seen in Figure 2.2, but also aligns with the notion of persistent divergence if a country enters a "violence trap"—a repeated cycle of peace and violence. Figure 2.4 demonstrates this more persistent divergence, which could explain the enduringly high poverty rates in countries such as Afghanistan, the Democratic Republic of Congo, or Somalia, where such cycles are prevalent.

In other words, violent armed conflict could play a very important part in keeping countries from fighting poverty effectively. However, when interpreting these relationships, it should be kept in mind that it is likely that armed conflict is only part of a larger relationship between weak institutions, a weak economy and a weak state with causal links between all four.

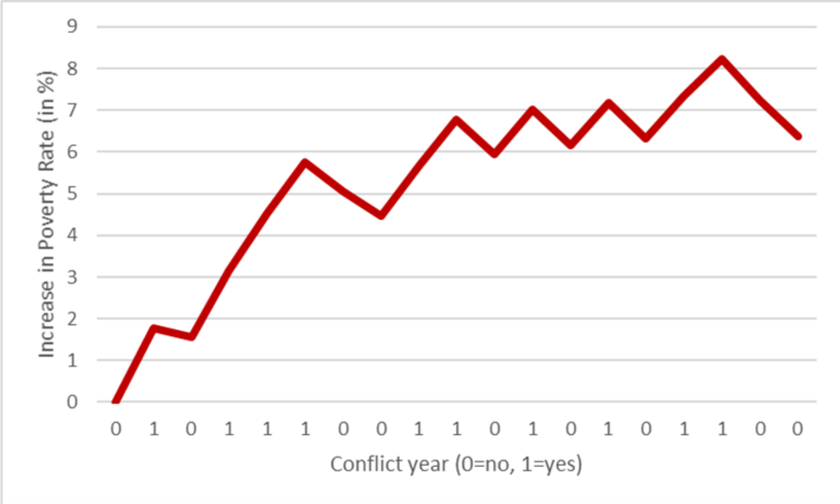
What our empirical analysis has demonstrated is that poverty reduction often failed in areas which were plagued by conflict. It is less clear whether reducing armed conflict will automatically lead to a reduction in poverty, i.e. it is not certain whether peace is a sufficient condition for poverty reduction. But it has to be kept in mind that the much better identified micro literature finds dramatic effects of conflict exposure on physical and mental health, education levels and asset evaluations. If the size of the effect found in these case studies is any guide then it is entirely realistic that intense conflict as we study it here is responsible for a population suffering from reduced human capital and low investments for years after the violence has stopped. So, whereas the macro results by themselves do not constitute conclusive evidence they strongly suggest a causal relationship in the context of the micro literature on the effect of violence on human capital.

Figure 2.3: Simulation of Dynamics in Poverty with a Single Conflict Episode




Notes: Calculations are based on Equations 2.1 and 2.2 in the text. As parameter estimates we use $\delta$ equal to 0.88 and the estimate $\beta$ from Column (3) of Table 2.2.

Figure 2.4: Simulation of Dynamics in Poverty in a Conflict Trap

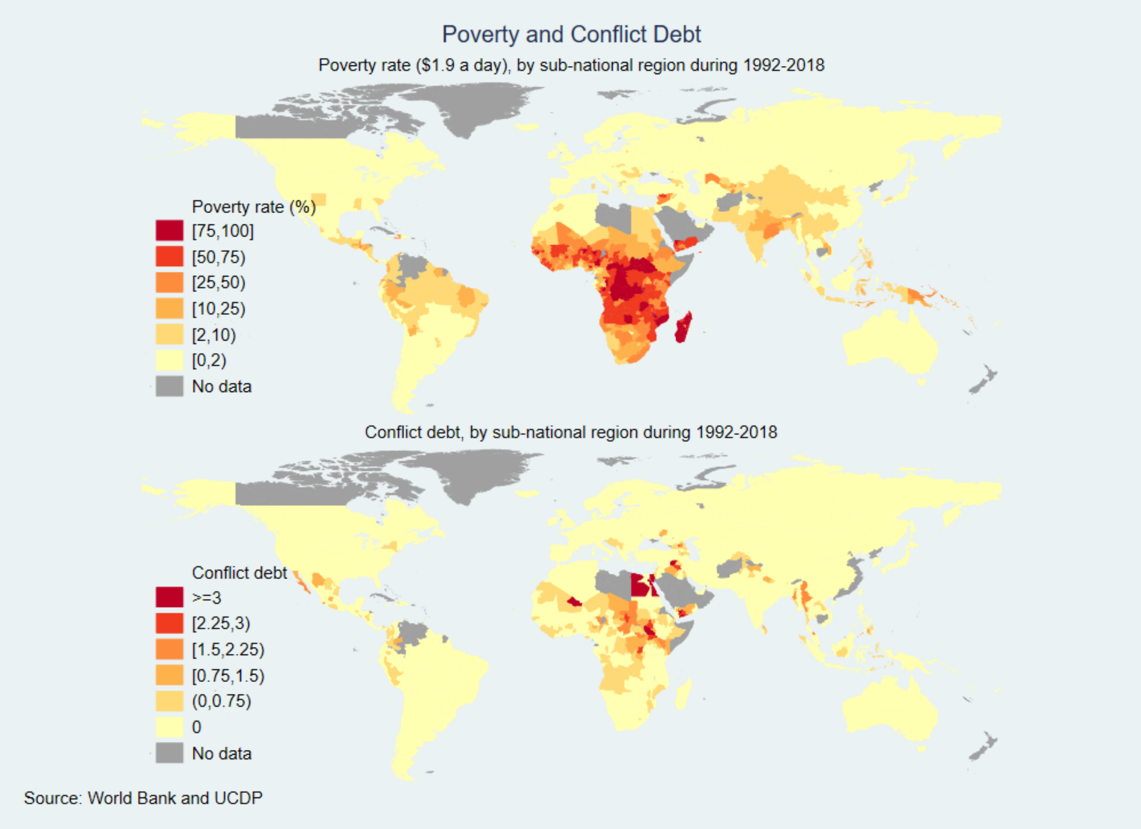



Notes: Calculations are based on Equations 2.1 and 2.2 in the text. As parameter estimates we use $\delta$ equal to 0.88 and the estimate $\beta$ from Column (3) of Table 2.2.

### 2.4.3 Regional Results

We now turn to the regional data, enabling us to investigate the empirical relationship between conflict history and poverty at the subnational level. In Figure 2.5, we present both the average poverty levels and the average conflict debt for the same year across all Admin-Level-1 regions included in the dataset. Notably, the conflict history has been matched to the corresponding region and year of the poverty data's availability.

Figure 2.5: Poverty and Conflict History in Admin-Level-1 Regions




Sources: PovcalNet by World Bank, and UCDP. Notes: Conflict debt is the discounted history of conflict years where a conflict year is defined as by more than 0.08 fatalities per 1,000 population.

The main insight from this data is that poverty is indeed often concentrated in regions which also have a high conflict debt. In Table 2.4 we confirm this with regressions similar to the country level but now weighted by population.[4] Strikingly, the coefficients we find in Columns (1) and (2) of Table 2.4 are very similar in magnitude as the cross-country finding in Column (1) of Table 2.2. An important question is, however, whether this is driven by the within-country or between country levels of poverty and violence. In Column (3) we show that the coefficient is extremely similar to the coefficients found in Column (3) despite the fact that the variation is now within country across regions instead of within country across time. But we also find that the country level conflict debt is strongly associated with regional poverty in Columns (4) and (5). When comparing the models in Columns (6) and (7), the country-level debt produces stronger results; however, it remains challenging to entirely discount the influence of local effects.[5]

This suggests both regional and country conflict debt plays a role for poverty but

[4]The main idea of weighting by population is that we want to reduce the level of noise coming from regions with only few survey respondents. The coefficients we find without population weighting are similar but precision falls dramatically.

[5]In Table 2.10 in the Appendix, we demonstrate that the results remain robust even when excluding countries for which only aggregate country-level observations are available.

that there are important regional spill-overs at work which affect the regional result. One obvious channel through which this will be the case is internal displacement which is affecting a large number of people.[6] Displacement will lead to regional conflict debt and poverty measures to diverge if the channel is the loss of individual human capital which is displaced at the same time at which it is harmed. Another possible channel is that risk perceptions change for an entire country if a region experiences violence. Outside investors will then not invest in any region of a country. This might also explain why subnational studies sometimes find less strong effects of violence on the economy when compared to macro studies.

Table 2.4: Poverty and Conflict Debt: Sub-country Level (Weighted Regressions)

| | (1)<br>Poverty rate | (2)<br>Poverty rate | (3)<br>Poverty rate | (4)<br>Poverty rate | (5)<br>Poverty rate | (6)<br>Poverty rate | (7)<br>Poverty rate |
|---|---|---|---|---|---|---|---|
| Conflict debt | 11.27*** | 4.672** | 2.279 | | | 7.005** | 2.581 |
| | (3.252) | (1.880) | (1.484) | | | (3.557) | (2.121) |
| Country conflict debt | | | | 16.01*** | 7.785*** | 12.03*** | 6.374** |
| | | | | (3.225) | (2.240) | (3.469) | (2.519) |
| Country fixed effects | No | No | Yes | No | No | No | No |
| Continent fixed effects | No | Yes | No | No | Yes | No | Yes |
| Adjusted R-squared | 0.0279 | 0.527 | 0.737 | 0.0352 | 0.530 | 0.0435 | 0.531 |
| Observations | 1885 | 1885 | 1885 | 1885 | 1885 | 1885 | 1885 |

Notes: Robust standard errors in parentheses. * p<0.1, ** p<0.05, *** p<0.01. The dependent variable is poverty rate (in percent). Conflict debt is the discounted history of conflict years where a conflict year is defined as by more than 0.08 fatalities per 1,000 population. Country conflict debt is computed in the same manner at the country level. The regressions weighted by population are on the cross-sectional data of sub-country regions between 1992-2018 whenever sub-regional poverty rates are available.

In Table 2.5 we use both the UCDP and the ACLED data to look into the relationship between conflict history and the poverty rate at the regional level but within the continent of Africa. As before we use population weights to reduce measurement error. Columns (1) and (2) show results for UCDP and Columns (3) and (4) show results for ACLED. All results are for the African continent only and are remarkably robust between the two conflict datasets. Most importantly, the results are now robust to country fixed effects in Columns (2) and (4), i.e. they are significant within countries. This provides some evidence for a local mechanism being at play. It is again remarkable how the coefficient on conflict debt is robust across tables with magnitudes being very similar in Tables 2.2, 2.4, and 2.5.

[6] According to the UNHCR some 41.3 million people were internally displaced due to armed conflict at the end of 2018.

Table 2.5: Poverty and Conflict Debt at Sub-national Regions in Africa: UCDP and ACLED data

| | (1) Poverty rate | (2) Poverty rate | (3) Poverty rate | (4) Poverty rate |
|---|---|---|---|---|
| Conflict debt | 14.13*** | 5.705* | 9.974*** | 5.295** |
| | (3.670) | (3.055) | (3.182) | (2.140) |
| Violence data | UCDP | UCDP | ACLED | ACLED |
| Country fixed effects | No | Yes | No | Yes |
| Adjusted R-squared | 0.0425 | 0.602 | 0.0322 | 0.604 |
| Observations | 618 | 618 | 618 | 618 |

Notes: Robust standard errors in parentheses. * p<0.1, ** p<0.05, *** p<0.01. The dependent variable is poverty rate (in percent). Conflict debt is the discounted history of conflict years where a conflict year is defined as by more than 0.08 fatalities per 1,000 population. The regressions are on the cross-sectional data of sub-country regions in Africa between 2005-2018 whenever sub-regional poverty rates are available.

### 2.4.4 Evidence from within Uganda

To explore the dynamics at the regional level, we examine conflict history and poverty rates through a case study of Uganda, utilizing comparable regional poverty data across two time points: 2012 and 2016. This period follows the most intense fighting in Uganda, necessitating the use of detailed ACLED data to capture any armed violence during this time. However, our primary focus is on regional recovery trends. If our proposed framework is correct, we would expect to observe a reversion to the mean in the regions that experienced the highest levels of violence, with these regions now catching up in terms of poverty reduction relative to the rest of the country. The data indeed confirm this pattern, supporting the validity of our model.

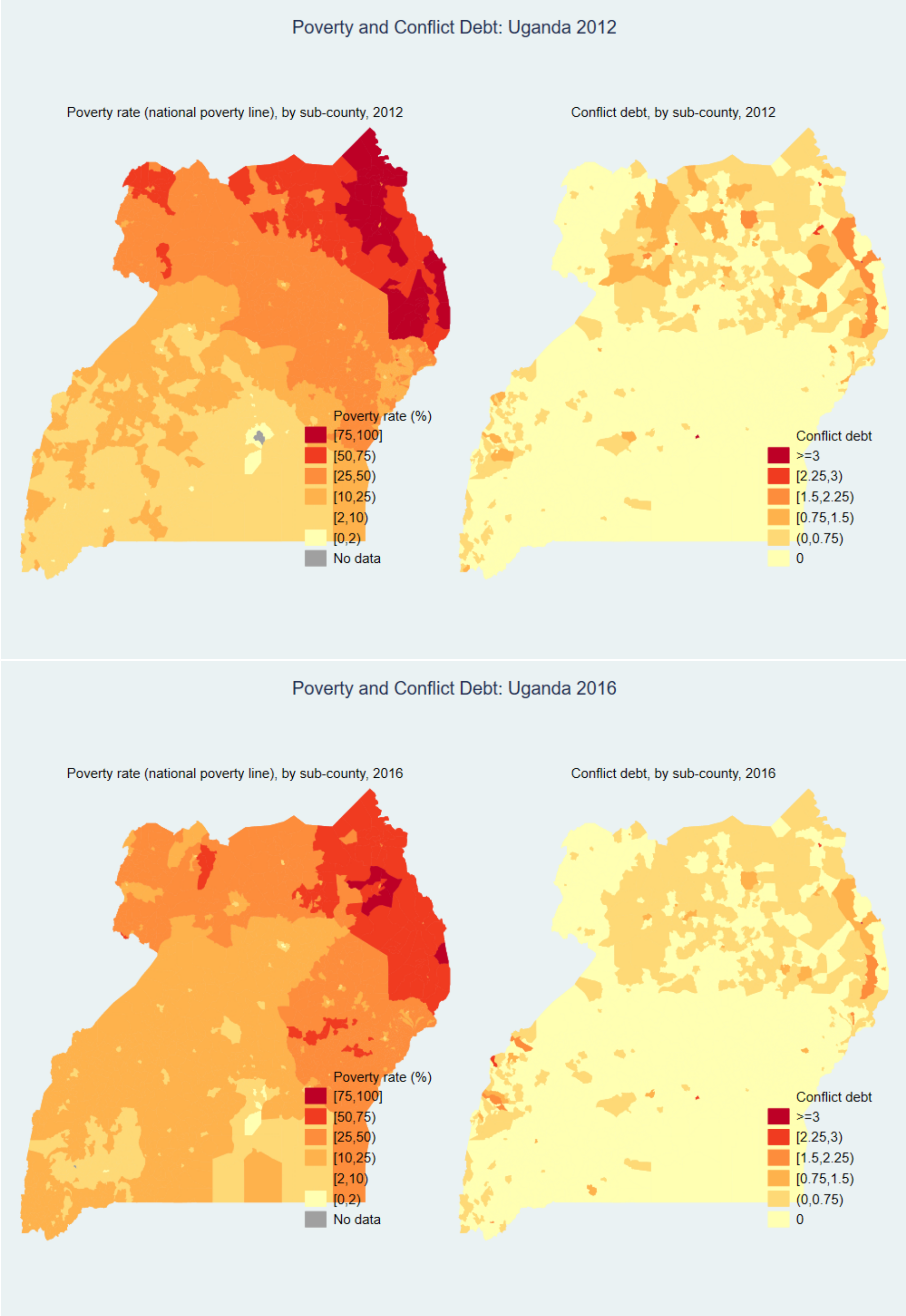
Figure 2.6 shows a strong correlation between conflict debt and poverty within Uganda. In particularly, the formerly violent North is also poorer. Second, as conflict debt is reduced over time, so is poverty. This raises the question of whether regions that were most heavily affected by poverty are catching up with the rest of the country. We test this more specifically in Table 2.6 . Here we first look at the conflict debt at the sub-county level. We control for regional fixed effects and a time dummy for 2016 to capture general trends in poverty. There is a statistically and economically significant relationship between conflict history and the poverty rate at the regional level. Importantly, this relationship holds controlling for sub-country fixed effect which rules out long-term fixed characteristics driving the relationship. This relationship between conflict debt and poverty also holds if we aggregate the violence data up to the county and even the district level. This is in line with the findings in Table 2.4 which suggest spill-overs. In Column (4) we conduct a horse-race and find that there is relevant catch up at the county level even when controlling for the district level history. However, we also find that debt at the sub-country is not significant, i.e. the much smaller areas for which we have data, are probably not the right level of aggregation to understand the impact of conflict on poverty.

Table 2.6: Poverty and Conflict History in Uganda 2012 and 2016

| | (1) Poverty rate | (2) Poverty rate | (3) Poverty rate | (4) Poverty rate |
|---|---|---|---|---|
| Conflict debt | 10.26*** | | | 1.928 |
| | (1.606) | | | (1.515) |
| County conflict debt | | 13.02*** | | 7.556*** |
| | | (0.975) | | (1.103) |
| District conflict debt | | | 11.11*** | 5.345*** |
| | | | (0.730) | (0.592) |
| Sub-county fixed effects | Yes | Yes | Yes | Yes |
| Year fixed effects | Yes | Yes | Yes | Yes |
| Adjusted R-squared | 0.121 | 0.206 | 0.200 | 0.215 |
| Observations | 3026 | 3026 | 3026 | 3026 |

Notes: The dependent variable is poverty rate (in percent). Conflict debt is the discounted history of conflict years where a conflict year is defined as by more than 0.08 fatalities per 1000 population. We use average population to avoid time variation in the denominator of the indicator. County and district conflict debts are computed in the same manner at higher administrative units, respectively. Regressions are on the balanced panel dataset in 2012 and 2016. Robust standard errors in parentheses. * p<0.1, ** p<0.05, *** p<0.01

Figure 2.6: Relative Poverty Reduction in Uganda's Formerly Violent North, 2012 and 2016




Sources: PovcalNet by World Bank, and ACLED. Notes: Conflict debt is the discounted history of conflict years where a conflict year is defined as by more than 0.08 fatalities per 1,000 population.

## 2.5 Summary

In this short note we derived a measure which is motivated by the micro literature, theoretical considerations and makes sense of the dynamic relationship between poverty and armed conflict. We call the derived measure *conflict debt* as it behaves like debt which accumulates during conflict and only fades in its impact over time. Conflict debt is strongly associated with poverty rates at the country level and the subnational level.

A novel aspect of our model is the recognition that armed conflict can both hinder poverty reduction and, once concluded, allow for significant catch-up effects. This is a critical finding, as it provides a connection to subnational studies, such as those on Vietnam by Miguel and Roland (2011), who found no long-term relationship between past violence (e.g., bombings) and various poverty-related outcomes. Our results suggest that such catch-up effects may explain this lack of relationship, as after a decade of peace, one might expect only a mild correlation between the intensity of past conflict and current poverty levels.

Whereas each single result in this note should not be interpreted as causal evidence, the consistent association between conflict debt and poverty across different datasets and levels of aggregation strongly suggests that we are picking up a causal link from conflict to poverty. However, it is important here to stress the micro literature which is better identified and proposes a channel running through human capital accumulation and low expectations. In line with this we find that the deprivation from education and living standards are strongly associated with conflict debt at the macro level.

If we are conservative and take the result in Column (3) of Table 2.2 as a guide, the most conflict-ridden countries and regions would have 5-10 percentage points lower poverty rates without their conflict debt. However, if one takes the simple cross-country and regional correlations as a guide the implied poverty reduction from eradicating conflict debt could be more than twice as high.

Our findings offer further evidence for the critical importance of preventing and deescalating conflict. While approaches to achieving this differ, recent evidence highlights the effectiveness of power sharing and building institutional legitimacy as valuable strategies. Although such measures may not always be feasible at the national level, there is evidence that local power-sharing arrangements can effectively reduce violence at the community level (Fetzer and Kyburz, 2018; Mueller and Rohner, 2018; Bormann et al., 2019).

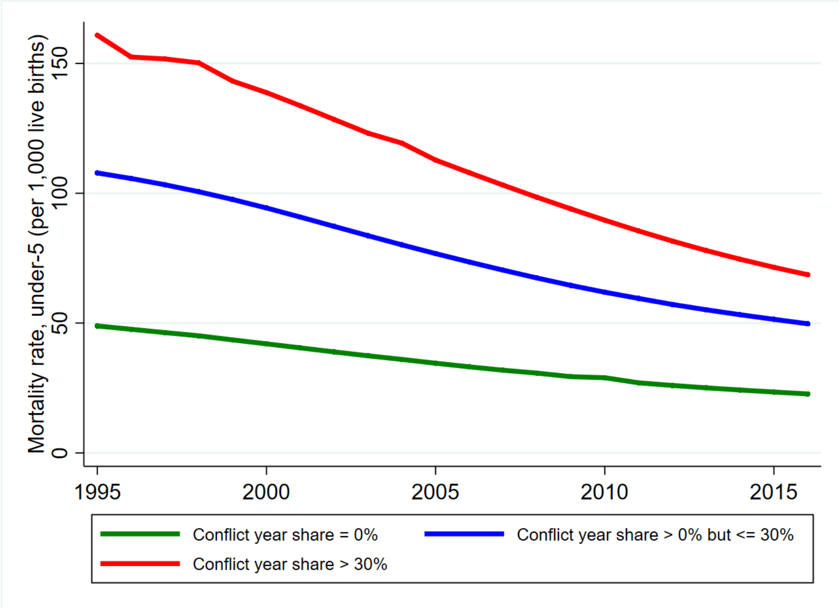
Data concerns persist, particularly in conflict-affected areas. As noted by Corral et al. (2020), poverty measurement is most challenging in the regions and periods experiencing violence. Uganda may exemplify this issue, as poverty measurements were conducted after the most intense violence had subsided. This underscores the pressing need for data from countries that have only recently emerged from conflict to accurately assess the impact of armed conflict on livelihoods.

# Appendix for Chapter 2

## B1 Figures

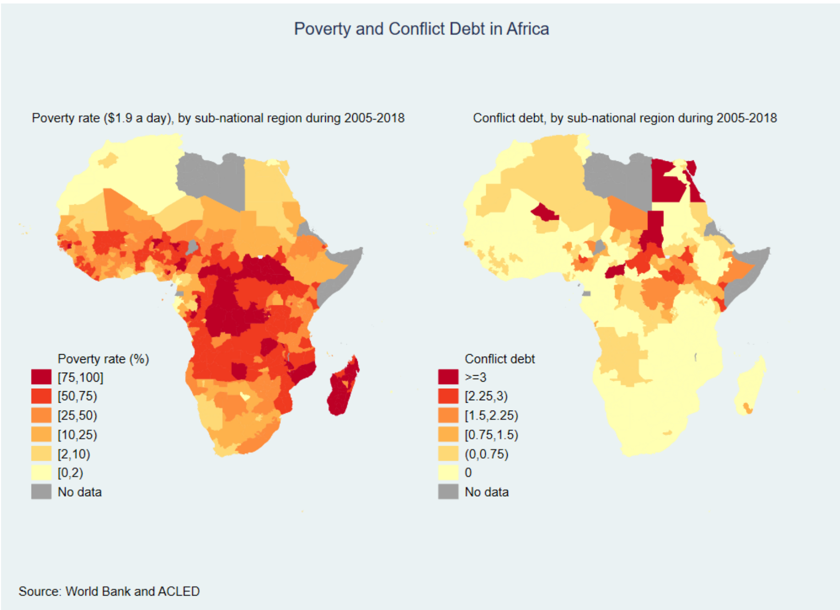
Figure 2.7 shows the relationship between child mortality and conflict history. We find a somewhat different pattern if we look at child mortality instead of poverty rates (shown in the main text). Here the catch-up dynamics in countries affected by conflict are overwhelming. These results are in line with the findings on the nutrition dimension in Table 2.9. Child mortality catches up very quickly after conflict ends and there is therefore no significant relationship between mortality and conflict debt in the long run.

Figure 2.7: Child Mortality and Conflict History




Sources: World Bank, and UCDP. Notes: A conflict year is defined by a year with 0.08 fatalities per 1,000 population. We use average population to avoid time variation in the denominator of the indicator. Average population is computed from available years in the World Bank data. Conflict year share is a share of conflict year during 1990-2019

Figure 2.8: Poverty and Conflict History at Admin1 Level in Africa using ACLED data




Sources: PovcalNet by World Bank, and ACLED. Notes: Conflict debt is the discounted history of conflict years where a conflict year is defined as by more than 0.08 fatalities per 1,000 population.

# B2 Tables

Table 2.7: Poverty Rate and Conflict: Cross-country Level

| | (1) Poverty rate | (2) Poverty rate | (3) Poverty rate | (4) Poverty rate |
|---|---|---|---|---|
| Conflict debt | 9.961*** | 9.831*** | 1.780* | 1.864* |
| | (1.295) | (1.271) | (1.060) | (1.062) |
| Conflict year | -8.752* | -10.18** | -0.0595 | -1.843 |
| | (4.803) | (4.698) | (1.254) | (1.417) |
| Country fixed effects | No | No | Yes | Yes |
| Time fixed effects | No | Yes | Yes | No |
| Continent-time fixed effects | No | No | No | Yes |
| Adjusted R-squared | 0.0520 | 0.0824 | 0.276 | 0.422 |
| Observations | 3238 | 3238 | 3238 | 3238 |

Notes: Robust standard errors clustered at the country level in parentheses. * p<0.1, ** p<0.05, *** p<0.01 The dependent variable is poverty rate (in percent), interpolated linearly to replace missing values between two data points. A conflict year is defined by a year with 0.08 fatalities per 1,000 population. We use average population to avoid time variation in the denominator of the indicator. Average population is computed from available years in the World Bank data. Conflict debt is the discounted history of conflict years. Continent-time fixed effects are a dummy for every year on every continent. Regressions are on unbalanced cross-country sample during 1989-2018.

Table 2.8: Poverty Rate and Debt Cut-off, Cross-country Level

| | (1) Poverty rate | (2) Poverty rate |
|---|---|---|
| Minor debt burden | -1.941 | 1.624 |
| | (1.287) | (1.276) |
| Some debt burden | 13.75*** | 14.47*** |
| | (1.694) | (1.757) |
| Heavy debt burden | 21.13*** | 20.43*** |
| | (1.886) | (1.872) |
| Country fixed effects | No | No |
| Time fixed effects | No | Yes |
| Adjusted R-squared | 0.0774 | 0.108 |
| Observations | 3238 | 3238 |

Notes: Robust standard errors clustered at the country level in parentheses. * p<0.1, ** p<0.05, *** p<0.01. The dependent variable is poverty rate (in percent), interpolated linearly to replace missing values between two data points. Discounted cumulative violent years is computed from dummy of violent year + 0.88*Lagged 1 of discounted cumulative year. The minor debt is 0 up to .19, some is up to .68 and heavy is over 0.68 Regressions are on unbalanced cross-country sample during 1990-2017.

Table 2.9: Multidimensional Poverty and Conflict Debt, Cross-country Level

| | MPI headcount ratio (%) | MPI headcount ratio (%) | Deprived in nutrition (%) | Deprived in child mortality (%) | Deprived in years of schooling (%) | Deprived in school attendance (%) |
|---|---|---|---|---|---|---|
| Conflict debt | 9.342*** | 4.146*** | 1.514 | 1.088 | 2.372** | 3.210*** |
| | (2.189) | (1.133) | (0.919) | (0.791) | (0.919) | (1.033) |
| Time fixed effects | No | Yes | Yes | Yes | Yes | Yes |
| Adjusted R-squared | 0.0703 | 0.682 | 0.496 | 0.551 | 0.475 | 0.470 |
| Observations | 293 | 293 | 267 | 272 | 293 | 272 |
| | Deprived in cooking fuel (%) | Deprived in sanitation (%) | Deprived in drinking water (%) | Deprived in electricity (%) | Deprived in housing (%) | Deprived in assets (%) |
| Conflict debt | 3.700*** | 3.328*** | 4.563*** | 3.699*** | 4.287** | 4.376*** |
| | (1.234) | (1.096) | (0.908) | (1.165) | (1.732) | (1.360) |
| Time fixed effects | Yes | Yes | Yes | Yes | Yes | Yes |
| Adjusted R-squared | 0.675 | 0.670 | 0.644 | 0.696 | 0.531 | 0.558 |
| Observations | 284 | 293 | 293 | 288 | 284 | 292 |

Notes: Robust standard errors clustered at the country level in parentheses. * p<0.1, ** p<0.05, *** p<0.01. The dependent variable is poverty rate (in percent). Conflict debt is the discounted history of conflict years where a conflict year is defined as by more than 0.08 fatalities per 1,000 population. Country conflict debt is computed in the same manner at the country level. The regressions weighted by population are on the cross-sectional data of sub-country regions between 1992-2018 whenever sub-regional poverty rates are available.

Table 2.10: Poverty and Conflict Debt: Sub-country Level without Country Observations (Weighted Regressions)

| | (1) Poverty rate | (2) Poverty rate | (3) Poverty rate | (4) Poverty rate | (5) Poverty rate | (6) Poverty rate | (7) Poverty rate |
|---|---|---|---|---|---|---|---|
| Conflict debt | 10.98*** | 4.522** | 2.279 | | | 6.844* | 2.519 |
| | (3.243) | (1.899) | (1.475) | | | (3.547) | (2.130) |
| Country conflict debt | | | | 15.59*** | 7.549*** | 11.71*** | 6.178** |
| | | | | (3.222) | (2.262) | (3.468) | (2.532) |
| Country fixed effects | No | No | Yes | No | No | No | No |
| Continent fixed effects | No | Yes | No | No | Yes | No | Yes |
| Adjusted R-squared | 0.0268 | 0.522 | 0.736 | 0.0338 | 0.525 | 0.0418 | 0.525 |
| Observations | 1862 | 1862 | 1862 | 1862 | 1862 | 1862 | 1862 |

Notes: Robust standard errors in parentheses. * p<0.1, ** p<0.05, *** p<0.01. The dependent variable is poverty rate (in percent). Conflict debt is the discounted history of conflict years where a conflict year is defined as by more than 0.08 fatalities per 1,000 population. Country conflict debt is computed in the same manner at the country level. The regressions weighted by population are on the cross-sectional data of sub-country regions between 1992-2018 whenever sub-regional poverty rates are available.

Table 2.11: Poverty and Conflict History in Uganda 2012 and 2016 (weighted regressions)

| | (1)<br>Poverty rate | (2)<br>Poverty rate | (3)<br>Poverty rate | (4)<br>Poverty rate |
|---|---|---|---|---|
| Conflict debt | 10.93*** | | | 0.798 |
| | (1.953) | | | (2.016) |
| County conflict debt | | 14.34*** | | 7.959*** |
| | | (1.395) | | (1.775) |
| District conflict debt | | | 12.48*** | 6.381*** |
| | | | (0.850) | (1.193) |
| Sub-county fixed effects | Yes | Yes | Yes | Yes |
| Year fixed effects | Yes | Yes | Yes | Yes |
| Adjusted R-squared | 0.131 | 0.199 | 0.196 | 0.206 |
| Observations | 3026 | 3026 | 3026 | 3026 |

Notes: Robust standard errors in parentheses. * p<0.1, ** p<0.05, *** p<0.01. The dependent variable is poverty rate (in percent). Conflict debt is the discounted history of conflict years where a conflict year is defined as by more than 0.08 fatalities per 1000 population. We use average population to avoid time variation in the denominator of the indicator. County and district conflict debts are computed in the same manner at higher administrative units, respectively. Regressions are on the balanced panel dataset in 2012 and 2016.

# Chapter 3

# Foreign Aid for Natural Resources

## 3.1 Introduction

North-to-south development assistance is characterized by the transfer of resources from wealthier and more democratic countries to lower-income and less democratic nations. Another feature of this assistance is the exchange between resource-importing and resource-exporting countries. In 2010, countries that were actively providing aid imported an average of 38.9 million tonnes of fossil fuels, while recipient countries exported almost 4 million tonnes. This chapter examines the motivation behind this type of aid, specifically how the abundance of resources in developing countries attracts financial support from donor countries that have a high demand for those resources.

As developing countries continue to experience rapid growth, their demands for natural resources, particularly fossil fuels, also increase. According to projections from the U.S. Energy Information Administration (EIA), non-OECD countries are expected to see a 70 percent increase in energy consumption between 2018 and 2050, compared to a 15 percent increase in OECD countries. A majority of this consumption growth is expected to come from Asian countries (EIA, 2019). Some previous studies have suggested that financial support from countries like China and India may be intended to gain access to natural resources in recipient countries (Dreher et al., 2018; Fuchs and Vadlamannati, 2013). Therefore, some policymakers have expressed concern that this type of foreign aid may negatively impact the developmental impact that aid is supposed to provide (Lum et al., 2009). Prior research has both supported and denied these accusations.[1]

However, it should be noted that the aid motivations of developed nations might not

[1] Studies that support the arguments are Li (2017), Brazys et al. (2017), Hernandez (2017), and Gehring et al. (2019); those rejected are Hackenesch (2015), Dreher and Fuchs (2015), and Dreher et al. (2018); Dreher et al. (2011) detail both similar and distinct aspects of aid from old and new donors.

be purely altruistic. Japan's foreign aid policy only began to respond to international pressures in the early 1990s, but before that, the growing industrial country used its aid programs to secure oil shipments from China and the Middle East. More recent Japanese aid programs in Myanmar, following the country's political reforms in 2010, can also be seen as driven by the desire for resources (Reilly, 2013). The UK government has also been criticized for allocating more funds to fossil fuel projects in resource-rich developing countries like Kuwait, Oman, and Brazil, despite its commitments under the Paris Agreement (Watt, 2019). Similarly, South Korea has shifted its aid destination from Asian countries to Africa since 2010, with a focus on oil-rich countries like Angola, Ghana, and Egypt (Marx and Soares, 2013). A study by Couharde et al. (2020) finds that the allocation of aid by G7 countries is strongly correlated with the oil endowment in recipient countries and the oil dependence of donor countries.[2]

Preliminary evidence suggests that foreign aid may be used as a means of gaining access to resources. Figure 3.1 illustrates the amount of official development assistance (ODA) committed by donors in the OECD database from 1990 to 2020, in comparison to gas and oil prices. The figure shows that, in general, aid commitments tend to follow the value of fossil fuels throughout the period.

Additionally, the majority of aid commitments come from resource-importing donors. Figure 3.2 illustrates the average aid recipients and average aid commitments per year by donor, compared to their average net fossil fuel imports. The positive association suggests that donors with high resource demand commit more aid contracts and larger amounts of aid. On the other hand, the number of aid contracts and the amount of aid commitments from resource-exporting donors are also positively correlated with the demand for fossil fuel imports in Figure 3.2. This figure suggests that resource-rich donors, located on the left side of the zero line, tend to contribute more aid, leading to a V-shaped relationship. Australia, Canada, and Norway are among the leading donors in this category.

This chapter examines the motivations behind foreign aid and contrasts two motivations related to resource values: altruism and resource-access motivation. It introduces a framework to test these hypotheses, making it one of the first economic studies, if not the first, to model resource motivation. The proposed theory suggests that aid motivated by resource access behaves differently from purely altruistic aid in three specific ways: in response to natural resource values, the recipient's income, and aid from other donors. The study's results, utilizing exogenous variation from global resource prices, indicate that resource-access motivation significantly influences the allocation of development aid among donors with high demand for natural resources, particularly as resource valuations

[2] To the best of my knowledge, this is the only economic research that has studied the "resource-access" motivation of foreign aid among a limited group of traditional donors.

in the recipient country increase. Further investigation into the heterogeneity among resource-importing donors reveals that aid from traditionally generous donors—such as France, Germany, Japan, South Korea, Sweden, the United Kingdom, and the United States—is generally positively correlated with the natural resource valuations in their recipient countries, demonstrating a strong indication of resource-access motivations.

The chapter comprises seven sections. The subsequent section provides a literature review of previous studies relevant to the research. Section 3.3 presents a formal theory that focuses on the interaction between two primary motivations: altruistic and resource-access motivations. Section 3.4 details the data, measurements, and empirical specifications used in the chapter. The findings are presented in Section 3.5. In the following section, we further investigate heterogeneous resource motivations among donors. The chapter concludes in Section 3.7 with a summary of the findings.

Figure 3.1: Foreign aid and resource prices

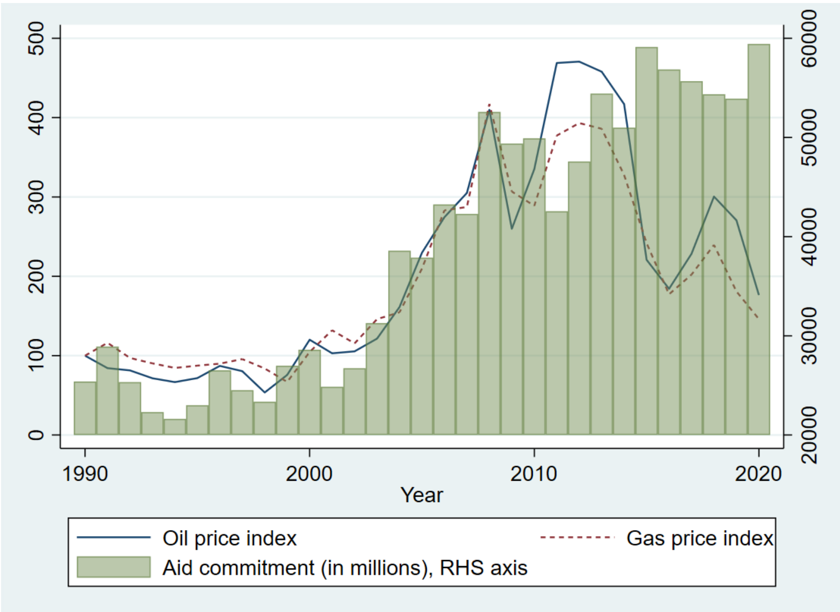
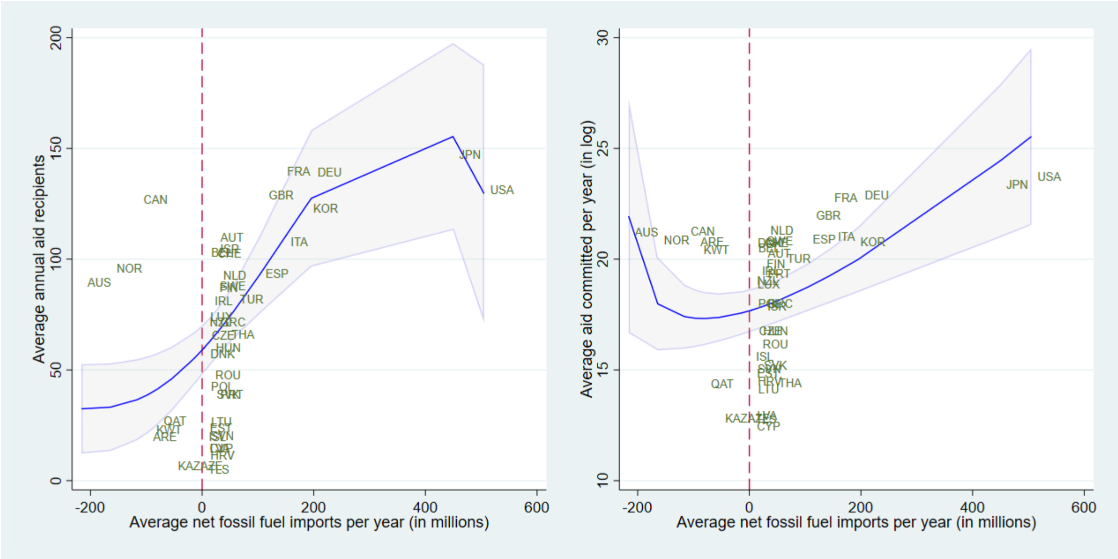


Figure 3.2: Aid commitments and resource imports by donor, 1990-2020




## 3.2 Related Literature

This chapter contributes to three distinct areas of economic research. The first area concerns the motivation behind foreign aid in general. The second area specifically explores links between natural resources and foreign aid. The third area pertains to the effectiveness of aid in promoting development, as the study provides an explanation for the fact that some aid is not targeted towards development, but instead towards quid pro quo activities, such as securing access to natural resources.

### 3.2.1 Motivations for Foreign Aid: Altruism and Self-Interest

In the literature on aid motivations, researchers often distinguish between two broad categories of aid: altruistic aid and aid driven by donors' interests. Altruistic aid aims to support the development of countries in need, whereas aid driven by donor interests reflects the strategic priorities of the donor countries. This self-interest can manifest in various ways, such as consistent voting patterns in international organizations, bilateral trade and investment relationships, and the colonial history of donor countries. Alesina and Dollar (2000) analyze bilateral aid flows from the OECD's Development Assistance Committee from 1970 to 1994, finding that colonial history and donors' political and strategic considerations are key factors in aid distribution. They noted donor behavior varies, with Nordic countries showing fewer political motives compared to France and Japan, where aid often aligns with colonial history, trade, and investment benefits. Hoeffler and Outram (2011) propose the "recipient merit" incentive, suggesting aid rewards countries with good policies like high economic growth and democratization. However, their data from 1980 to 2004 shows evidence of self-interest but limited support for the recipient merit hypothesis.

Other empirical studies have delved into the self-interest motives behind foreign aid. Kuziemko and Werker (2006) and Dreher et al. (2009) demonstrate that non-permanent membership in the United Nations Security Council (UNSC) is a driver for increased aid from the United States, the United Nations, and the World Bank. Furthermore, Pettersson and Johansson (2013) and Hühne et al. (2014) examine the relationship between foreign aid and trade flows between recipient and donor countries. Both studies find that bilateral aid is positively correlated with exports from both the recipient and the donor country. Specifically, Pettersson and Johansson (2013) provide evidence of aid-for-trade, finding that foreign aid is positively correlated with resource exports from recipient countries to donor countries.

The selectorate model proposed by de Mesquita and Smith (2009) suggest that policy concessions in aid-for-policy deals serve as public goods for donor citizens but public bads for recipients, particularly in countries with small selectorates that accept aid at lower costs. In contrast, our framework views resource access as a public good for both donor and recipient citizens, aligning their interests in aid-for-resource agreements.

### 3.2.2 Foreign Aid and Natural Resources

Resource abundance can be linked to foreign aid in several ways. We summarize prior research by identifying three potential mechanisms through which resource richness may influence the aid received: (i) recipients' demand for aid due to poverty and humanitarian crises, (ii) the supply of aid from resource-rich donors, and (iii) aid-for-resource agreements between resource-rich recipients and resource-poor donors. Each of these mechanisms will be discussed in turn.

The first channel involves recipients' demands for development, determined by changes in resource values. It can have an immediate impact when higher resource values boost recipient income, leading to a reduced demand for resources to boost the economy. This channel operates only if donors' intrinsic utility contains an altruistic motive that concerns the well-being of the population in developing countries.

This type of donor commits aid to promote development, and these altruistic transfers from north to south are plausible due to the decreasing marginal effect of aid on development. The same budget allocated to promote development in richer donor countries, such as poverty or inequality reduction, may be insufficient or even irrelevant when the problems relate to other causes and not a lack of budgets. In the literature, altruistic aid is promised to countries where living standards are subpar, and the recipient's income per capita represents the usual living standards. To link resource abundance to altruistic motivation, we consider natural resources owned by a recipient country to behave as windfall income that enhances financial resources for development by itself. Therefore, when the resource valuation increases, the altruistic assistance is weakened.

Within the first channel, another impact of resource values on demands for aid can be inverse if the natural resources lead to violent conflicts, a phenomenon known as the resource curse. A substantial body of literature indicates that an increase in resource values in countries with weak institutions can trigger violent conflict, as political actors vigorously compete for power, resulting in the resource curse (Collier and Hoeffler, 2004; Dube and Vargas, 2013; Berman et al., 2017; Andersen et al., 2022). As a result, countries facing such conflicts often receive more humanitarian aid from foreign sources to alleviate the negative consequences on their populations.

The second mechanism is unique to resource-rich donor countries, where the fluctuations in domestic resource values impact their aid provision. According to Werker et al. (2009), OPEC countries demonstrate a close relationship between their aid contributions, oil prices, and aid received by their lower-income Muslim allies. Furthermore, Figure 3.2 provides additional evidence supporting the aid-supply hypothesis, revealing that countries with higher net exporting resources tend to have higher aid commitments. Subsequent studies have employed oil prices as an exogenous instrument of aid from Gulf countries to investigate the impact of aid on political violence and regime survival (Ahmed, 2012; Ahmed and Werker, 2015).

The final channel involves resource access based on the dyadic characteristics of donors and recipients, suggesting that aid transfers from resource-importing donors to resource-rich recipients serve as additional payments to secure access to natural resources in recipient countries. Few economic studies offer rigorous analysis on this topic. Couharde et al. (2020) employ panel data from G7 donors and 82 recipients, using oil reserves as an explanatory variable, and find that increasing oil reserves lead to higher aid from oil-dependent donors. Dreher et al. (2018) provide detailed aid data from China, showing that concessional aid, which is less generous and more commercially motivated, is driven by resource-access interests. Fuchs and Vadlamannati (2013) examine whether financial assistance from India differs from that of traditional DAC donors. Using lagged resource extraction in recipient countries as an instrumental variable, they find no correlation between aid received and donor resource demand. In this chapter, we aim to develop a formal theory to understand resource motivation in relation to primary altruistic motivations, and we then use exogenous variations in global resource prices to test the hypotheses derived from the theory.

### 3.2.3 Aid Effectiveness

Another area of literature that has been extensively researched is the effectiveness of aid on economic development. However, there is no consensus on the impact of aid on economic development in the literature. For example, Collier and Dehn (2001) find that aid is most effective during difficult times in recipient countries, as it can prevent adverse effects and maintain growth more effectively. Some studies have provided further evidence on the conditional effectiveness of aid based on policies and political institutions (Burnside and Dollar, 2000; Collier and Dollar, 2002; Kaya and Kaya, 2020; Maruta et al., 2020).

A recent study by Dreher et al. (2021) using comprehensive Chinese data finds immediate growth impact of Chinese aid. Galiani et al. (2017) rely on a quasi-experiment setting of eligibility for aid from the International Development Association (IDA) and find a positive and sizable impact of aid on economic growth. However, there are also studies that assert that there are no growth effects from development aid (Easterly, 2003; Rajan and Subramanian, 2008). In fact, Asatullaeva et al. (2021) conducted a content analysis of the top 50 influential papers on the topic and found that 17 studies (38 percent) concluded that the effectiveness of aid on economic outcomes is conditional on policy, climate-related circumstances, or the type of aid. Meanwhile, 16 studies or 35 percent found that aid is ineffective, and 12 papers (27 percent) argued that aid is effective on economic development.

To explain opposing effects in the literature, Dreher et al. (2016) provide compelling evidence that the impact of aid on economic growth is undermined when the aid is politically motivated. Their results suggest that aid during the membership term of the United Nations Security Council (UNSC) has less impact on economic growth compared to the "normal" period. In this study, we present additional evidence supporting the strategic motivation related to natural resources and postulate that aid programs are designed not for development goals but for resource access. Hence, the development impacts of foreign assistance aid are not always anticipated.

The focus of this chapter is the implicit bargain between foreign aid donations and access to natural resources, a mechanism that is much less well-understood in the existing literature. This mechanism operates at the dyadic level, between natural resource-poor donors and resource-rich recipients. The theory outlined in the following section will examine this motivation within a theoretical framework and derive implications that enable the empirical distinction of this motivation from altruistic motives.

## 3.3 Theory of Aid for Natural Resources

This section introduces a static model to investigate the combined influence of altruistic and resource-access motivations on the decision to provide aid. We set up the aid commitment decision of a representative donor ($d$) to a representative recipient ($i$).

The donor's preference is characterized by two parts. The first is the development gain from committing aid to the recipient with the altruistic weight toward the recipient ($\lambda_{id}$). This term is represented by the utility of the recipient weighted by the donor's altruism. The second part is the utility from resource access, weighted by the donor's concern over resource security ($\delta_d$). The utility function of donor $d$ who proposes aid to recipient $i$, $A_{id}$, is

$$U(A_{id}; \lambda_{id}, \delta_d) = \lambda_{id}\ c(Y_i + A_{id} + A_{io} + \alpha_i R_i) + \delta_d r(A_{id}, A_{io}) R_i - A_{id}, \tag{3.1}$$

where $c(\cdot)$ is the recipient's utility function from consumption, such that $c'(\cdot) > 0, c''(\cdot) < 0$, $r(\cdot)$ is the contest function used to determine the share of resource values that donor $d$ receives compared to the other donors $o$, $\lambda_{id} \geq 0$ is the altruistic weight of donor $d$ to recipient $i$, $\delta_d \geq 0$ is the weight of donor's concern for resource access, $Y_i$ is the national income of recipient $i$, $A_{io}$ represents the other foreign aid received by country $i$ but not from donor $d$, $R_i$ is the value of resource values in country $i$, and $\alpha_i \in (0, 1)$ is the share of value utilized for the annual budget by resource-rich recipient $i$.

The initial component of the utility function represents the altruistic gain from the recipient's development. The utility function from consumption $c(\cdot)$ satisfies a set of standard assumptions, whereby the utility function increases as the national income, resource values, and the total foreign aid increase. It is important to note that this altruistic motivation incurs even if donor $d$ commits no assistance, but the (plausible) receiver can still enjoy their consumption. On the contrary, if the donor lacks concern for the recipient's welfare, the altruistic gain will be zero.

The second part of the utility is the gain from resource access. This motivation comprises the weighted gain from resource access ($\delta_d$), the overall resource value in the receiving country $i$ ($R_i$), and the share of resource value received by donor $d$ from the contest function. The contest function $r(A_{id}, A_{io})$ depends on donor's aid and other donors' aid competing for the share of resource values. We make some assumptions about the properties of this function.

Assumption 1. The contest function, $r(A_{id}, A_{io})$, exhibits the following.

1. $r(A_{id}, A_{io}) \in [0, 1]$, where $r(0, A_{io}) = 0$ and $r(A_{id}, 0) = 1$,
2. $r'_d > 0$, and $r''_d < 0$,
3. $r'_o < 0$, and $r''_o > 0$,

where $r'_d$, $r''_d$, $r'_o$, and $r''_o$ represent the first and second derivatives of the first and second arguments, respectively. Assumption 1.1 outlines the standard features of a contest function, ensuring that it is bounded between zero and one. Assumptions 1.2 and 1.3 specify that the function $r(\cdot)$ increases and is concave in $A_{id}$, and decreases and is convex in $A_{io}$.

The increase in resource values means that, with everything else held constant, the donor benefits more from the recipient if the donor contributes more assistance. The concavity in $A_{id}$ results from rent-seeking behavior in the contest function, in which donor $d$ and others $o$ compete for resource access. Holding all other efforts constant, a tighter commitment of aid results in a larger share of the total value, but at a diminishing rate.[3] This also applies to foreign aid from other donors, wherein it decreases the rate of contest value decrease as well, according to Assumption 1.3. In addition, the decreasing marginal return can arise from a notion of resource diversification. That is, the security gain from the resource supplied by country $i$ is diminishing once the values from this source is becoming more prominent to the donor.

One underlying assumption from the functional form of $\delta_d r(A_{id}, A_{io}) R_i$ is that the gain in resource access from an additional unit of aid increases when the resource value is higher. This assumption is rooted in the idea that providing aid to a scared-resource recipient would result in a marginal gain in securing resources that is lower than that of a abundant-resource recipient. This premise incentivizes donors to engage with resource-rich countries, as long as $\delta_d > 0$. It should be noted that in the event of $\delta_d = 0$, the utility gain from aid is solely derived from the altruistic component.

In this model, awarded natural resources are regarded as a distinct type of policy concession that generates positive externalities for the donor country. Resource-access aid constitutes an additional source of income, arising from the demand for resource security, alongside the commercial revenues paid by multinational oil companies. Aid recipients invariably benefit from such contracts, regardless of whether the donor's motives are purely altruistic or partially driven by resource acquisition. Consequently, a resource-rich, aid-receiving country consistently accepts aid offers from donors, and the recipient's optimization problem is therefore excluded from the analysis.

The utility function incorporates a basic monetary cost of aid itself as the disutility, reflecting the stylized fact that aid donors are typically high-income states and their budgets for this department constitute a small fraction of their income. Combining this fact with the assumption of decreasing marginal utility, any utility loss from committing aid is likely to be relatively small and constant, leading to the efficiency of altruistic aid transfer.

The final simplification made by the proposed utility function is that foreign aid from other donors is assumed to be given to the model, meaning there is no strategic feedback

[3] For example, the contest function developed by Tullock (2001), which is frequently used in the rent-seeking literature, exhibits increasing rent-seeking effort at a decreasing rate.

from other donors. This assumption may arise when the aid market is large enough and structured in a way that a single donor's decision cannot significantly influence the behavior of others. Another reason for this simplification is the absence of an appropriate notion of the cross-derivative between aid from donor $d$ and that from other donors $o$. The cross-derivative can be positive, indicating that more aid from other donors coordinates with a donor and intensifies the marginal return from the resources. Conversely, donors' aid may compete, undermining the marginal return of a single donor, resulting in a negative impact on the marginal value. We will later explore an empirical estimation to understand the cross-derivative relationship of aid among donors.

From the utility function 3.1, the first-order condition with respect to aid commitment is,

$$\lambda_{id}c'(Y_i + A_{id} + A_{io} + \alpha_i R_i) + \delta_d r'_d(A_{id}, A_{io})R_i = 1. \tag{3.2}$$

Define the optimal aid that satisfies Equation 3.2 as $\hat{A}(Y_i, R_i, A_{io}; \lambda_{id}, \delta_d)$, or $\hat{A}_{id}(Y_i, R_i, A_{io})$ in short. The first-order condition can be rewritten as

$$\lambda_{id}c'(Y_i + \hat{A}_{id}(Y_i, R_i, A_{io}) + A_{io} + \alpha_i R_i) + \delta_d r'_d(\hat{A}_{id}(Y_i, R_i, A_{io}), A_{io})R_i = 1, \tag{3.3}$$

which is in terms of the recipient's income, resource values, and aid from other donors. Differentiating Equation 3.3 with respect to $R_i$, $Y_i$, and $A_{io}$ on both sides of the equation provides three following results.

$$\frac{\partial \hat{A}_{id}(Y_i, R_i, A_{io})}{\partial R_i} = \frac{\alpha_i \lambda_{id} c'' + \delta_d r'_d}{-(\lambda_{id}c'' + \delta_d R_i r''_d)} \gtreqless 0. \tag{3.4}$$

$$\frac{\partial \hat{A}_{id}(Y_i, R_i, A_{io})}{\partial Y_i} = \frac{\lambda_{id} c''}{-(\lambda_{id}c'' + \delta_d R_i r''_d)} < 0, \tag{3.5}$$

$$\frac{\partial \hat{A}_{id}(Y_i, R_i, A_{io})}{\partial A_{io}} = \frac{\lambda_{id} c'' + \delta_d R_i r''_{do}}{-(\lambda_{id}c'' + \delta_d R_i r''_d)} \gtreqless 0. \tag{3.6}$$

Firstly, we compare the optimal responds between purely altruistic donors and ones with resource motivation. Condition 3.4, despite being ambiguous, can produce the following comparison

$$\left.\frac{\partial \hat{A}_{id}(Y_i, R_i, A_{io})}{\partial R_i}\right|_{\delta_d=0} = -\alpha < \left.\frac{\partial \hat{A}_{id}(Y_i, R_i, A_{io})}{\partial R_i}\right|_{\delta_d R_i>0}, \tag{3.7}$$

provided that Assumption 1.2 holds.[4] It suggests that in the absence of the resource-access motive (i.e., the pure-altruism model where $\delta_d = 0$), a rise in resource values in a

[4]Expression 3.7 holds true when $R_i > \frac{r'_d}{\alpha r''_d}$. This condition is met due to Assumption 1.2, where $\frac{r'_d}{r''_d} < 0$ and $R_i > 0$ as per the construction.

recipient country is expected to decrease the aid received from purely altruistic donors. On the other hand, donors with a resource motivation are less likely to react negatively to the resource windfall and may even increase aid if the demand for resources is high enough. We state this main result in the first proposition of how to detect the natural-resource motivation in the contrast to the altruistic one.

Proposition 1. *A donor motivated by resource interests will reduce foreign aid less than a purely altruistic donor, or may even increase it, as the value of the recipient country's resources rises.*

The next result of Expression 3.5 confirms a well-established finding in the literature: an increase in the recipient's national income decreases committed aid through altruistic utility. To explore heterogeneous reactions based on the donor's preference for resource access, we compare optimal aid responses from donors with and without a demand for resources ($\delta_d = 0$). Equation 3.5 implies that

$$\left.\frac{\partial \hat{A}_{id}(Y_i, R_i, A_{io})}{\partial Y_i}\right|_{\delta_d=0} < \left.\frac{\partial \hat{A}_{id}(Y_i, R_i, A_{io})}{\partial Y_i}\right|_{\delta_d R_i>0} \tag{3.8}$$

This means that donors with a demand for resources react less negative to economic development in resource-rich recipients than purely altruistic donors. Resource-concerned donors cannot completely withdraw their altruistic support when the demand for development is lower because foreign assistance is still needed to access natural resources. We state the first proposition from the result as follows.

Proposition 2. *A resource-motivated donor reduces aid to a resource-rich recipient less than a purely altruistic donor when the demand for development aid decreases.*

The last result of Expression 3.6 indicates ambiguous effects of others' aid on a single donor's decision. The impact can be positive if $r''_{do} > \frac{-\lambda_{id} c''}{\delta_d R_i}$, and non-positive otherwise. To explore the heterogeneous effects, Expression 3.6 posits that

$$\left.\frac{\partial \hat{A}_{id}(Y_i, R_i, A_{io})}{\partial A_{io}}\right|_{\delta_d=0} = -1. \tag{3.9}$$

In essence, Expression 3.9 hypothesizes the coordination of development aid among donors. With this notion, financial aid from other countries crowds out aid commitments from purely altruistic donors because this additional aid can similarly contribute to development outcomes in the recipient country without imposing a financial burden on the donor. However, the foreign aid literature may disagree with the substitutability but concur with the competitive aid, such that donors compete for relative development outcomes, resulting in aid fragmentation.[5] Still, this study maintains the simple coordinating altruistic aid model for simplicity, as it focuses on the resource-access motion.

The impact of other aid on a donor's resource-motivated aid depends on the functional form of the contest function, specifically whether $r''_{do}$ is greater than $r''_{d}$. If $r''_{do}$ is larger, indicating that aid efforts for resources among donors are competitive, we obtain

$$\left.\frac{\partial \hat{A}_{id}(Y_i, R_i, A_{io})}{\partial \hat{A}_{id}}\right|_{\delta_d=0} < \left.\frac{\partial \hat{A}_{id}(Y_i, R_i, A_{io})}{\partial A_{io}}\right|_{\delta_d R_i>0}, \text{ if } r''_d < r''_{do}.$$

The expression suggests that donors motivated by resources may react less negatively, or even positively, to increased contributions from other donors. In other words, they might maintain or increase their resource-motivated aid when other donors raise their aid levels. However, since we cannot definitively determine the contest function whether donors are competitive or cooperative when faced with larger aid contributions from others, we defer this question to the empirical analysis.

In summary, the theory identifies three channels through which resource-motivated donors behave differently from purely altruistic donors when providing financial aid to resource-abundant recipients. First, resource-motivated donors reduce aid less, or may even increase it, in response to the rising value of natural resources. Second, they exhibit a different response to environmental changes, showing less sensitivity to positive income shocks in the recipient country. Finally, resource-motivated donors are more likely to increase aid if competition for resources from other donors intensifies. In the empirical section, we conduct a test on the main finding regarding how donors respond differently to variations in natural resource value, while leaving hypotheses related to income and aid from other donors for future research.

## 3.4 From Theory to Empirical Tests

In this section, we first discuss the sources of data used in the estimation and how to construct measures to avoid endogeneity issues. Section 3.4.2 presents empirical specifications derived from the theoretical propositions.

### 3.4.1 Data and Measurements

The theory posits connections between recipient income, resource values of the recipient, other aid, and the foreign aid committed by a donor with preferences for altruism and resource security. This subsection introduces four principal variables relevant to the estimations and sample selection. The four variables are foreign aid variables from a

[5]For examples, see Steinwand (2015), Annen and Moers (2017), and Gehring et al. (2017).

donor and from other donors, natural resource values, the donor's level of concern regarding resource access, and the recipient's exogenous income shock. Lastly, the sample selection of donor-recipient pairs is discussed to test the hypotheses.

**Foreign aid.** The Official Development Assistance (ODA) commitment is sourced from the Creditor Reporting System (CRS) Aid Activities database, and it encompasses flows from 44 donors, including 34 members of the Organization for Economic Cooperation and Development (OECD) and 10 non-OECD countries, to low- and middle-income countries. This ODA commitment is provided by official aid agencies with the aim of promoting economic development and improving welfare in developing nations. It takes the form of concessional aid, consisting of soft loans and at least 25 percent grants (OECD, 2022). The bilateral aid between pairs of governments is calculated in constant value using the year 2020 as the base year.

The CRS database has a salient feature, which is the availability of information on the sector or purpose of aid projects. Consequently, we are able to distinguish between development aid *(DEV aid)* and emergency aid *(EMR aid)*.[6] The latter type of aid is suspected to emerge during violent conflicts, natural disasters, and humanitarian crises, and thus, reflects the indirect motivation of humanitarian aid driven by violence, which is not considered in the proposed theory. To prevent potential bias resulting from humanitarian motivation, we exclude *EMR aid* in most of the subsequent analyses.[7] In later estimation results, we compare the outcomes with different types of aid and demonstrate consistent estimates with our expectations.[8]

The development aid that is our main focus includes aid projects for:

1. Education ($Sector I1$),
2. Health ($I2$),
3. Population Policies/Programmes and Reproductive Health ($I3$),
4. Water Supply and Sanitation ($I4$),
5. Other Social Infrastructure and Services ($I6$),
6. Transport and Storage ($II1$),
7. Communications ($II2$),
8. Energy ($II3$),
9. Banking and Financial Services ($II4$),
10. Business and Other Services ($II5$),
11. Agriculture, Forestry, Fishing ($III1$),
12. Industry, Mining, Construction ($III2$),
13. Trade Policies and Regulations ($III3$),
14. General Environment Protection ($IV1$),
15. Other Multisector ($IV2$),

[6]"The emergency aid" consists of sectors in Government and Civil Society, Unallocated / Unspecified, Refugees in Donor Countries, General Budget Support, Development Food Assistance, Other Commodity Assistance, Action Relating to Debt, Emergency Response, Reconstruction Relief, and Rehabilitation, Disaster Prevention and Preparedness, Administrative Costs of Donors, Refugees in Donor Countries, and Unallocated / Unspecified.

[7]The distribution of dyadic aid contracts in the data, as illustrated in Figure 3.8, suggests that *DEV aid* is more commonly used than *EMR aid*. Estimated regressions in Appendix C2.2 show that emergency aid is more positively correlated with political violence in recipient countries than development aid.

[8]Further details on the dataset are provided in Appendix C1.1. Table 3.4 ranks the top ten recipient countries during 1990-2020, with India, China, and Indonesia being the top three, respectively.

We construct a panel dataset of donor-recipient pairs and calculate both a log-like transformation and percentiles relative to the entire distribution of dyadic aid flows, with the latter serving as the primary dependent variable. The aid percentiles are used instead of the absolute amounts for two reasons: the high proportion of zero cases and the extreme dispersion of the data. As shown in Figure 3.3, almost 25 percent of the data points have zero values, while the dispersion is extremely right-skewed, with the average absolute value slightly above the 80th percentile. Due to this long-tailed dispersion, large variations at the top of the distribution heavily influence the estimated coefficients at the mean value.

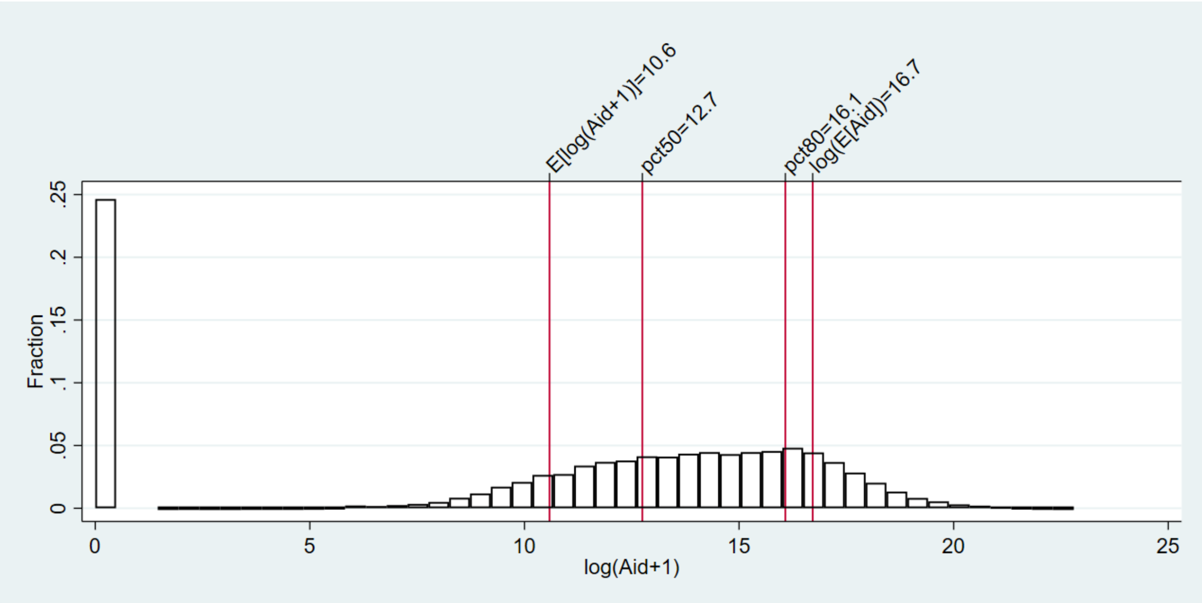
On the other hand, transforming the absolute variable into percentiles makes the distribution closer to a uniform distribution. The transformed outcome variable is more convenient for interpretation compared to the log-like transformation, where $log(Y + 1)$ doesn't provide a justified elasticity when the value starts at zero. Additionally, the underlying meaning of the percentile variable implies that the variation in committed aid matters when there is a change in ranking or its relative position within the entire distribution. This notion is particularly relevant in the context of aid commitment, where aid competition is prevalent in the literature and is suspected to be more prominent in the aid-for-resource context. The percentile transformation also helps prevent other variations, especially at the tail of the distribution, that could influence the estimation at the mean.[9]

And from the same dataset, we generate a control variable, which is the development aid from other donors to recipient $i$ ($A_{io}$). This variable is computed by subtracting the aid amount from donor $d$ from the overall foreign aid.[10]

[9] Couharde et al. (2020) select aid share allocated by donor to examine G7 donor allocations as the outcome variable. However, in this chapter, where the theory predicts distinct within-variation reactions among different types of donors—high- versus low-resource-demanding donors—relative aid to the entire distribution is more appropriate. For a discussion of rank transformation of the outcome variable and log-like transformation, see Chen and Roth (2023). Figure 3.9 in the Appendix displays example dyads that have shown the greatest improvement in their correlations by transforming the absolute aid values into percentiles.

[10] Figure 3.8 presents the distribution of the other aid compared to the aid received from donor $d$.

Figure 3.3: Distribution of development assistance




Note: The distribution of $log(Aid + 1)$ is presented instead of the distribution of the absolute values for better visibility.

**Natural-resource values.** We utilized data on oil and gas proven reserves and prices from the BP Statistical Review of World Energy to investigate the effect of resource values on aid allocation. Proven reserves refer to the volume that has a 90 percent or greater probability of being produced over the resource's lifetime (BP, 2021). To avoid endogeneity issues, we used the data on oil and gas reserves at the beginning of the study period in 1990. For countries established after 1990, we used the earliest available reserve data. We calculated the total value by multiplying the fixed reserve amount ($Q_{i,1990}$) by the corresponding annual prices ($p_t$). In addition, we created an indicator variable that takes a value of 1 if a country was a net exporter of fossil fuels from 1991 to 2019, using data from the UNEP IRP Global Material Flows Database. The formula we used to compute resource value is:

$$log(Resources_{it}) = \mathbb{1}(\sum_{t=1991}^{2019} \text{net fossil fuel export}_{it} > 0) \times log(Q_{i,1990} * p_t),$$

where $Q_{i,1990} * p_t = (\text{Oil reserves}_{i,1990} \times Price_{oil,t}) + (\text{Gas reserves}_{i,1990} \times Price_{gas,t})$.

Note that the imputed value represents the current valuation at time $t$ of fossil fuel reserves in 1990. Using fixed quantities in 1990 prevents the endogeneity problem where reserve quantity might be determined by other factors, including foreign assistance or other unobserved factors. This measure is assumed to be orthogonal to the error term in the estimation because the global market prices of resources are exogenous and cannot be determined by high-value countries.

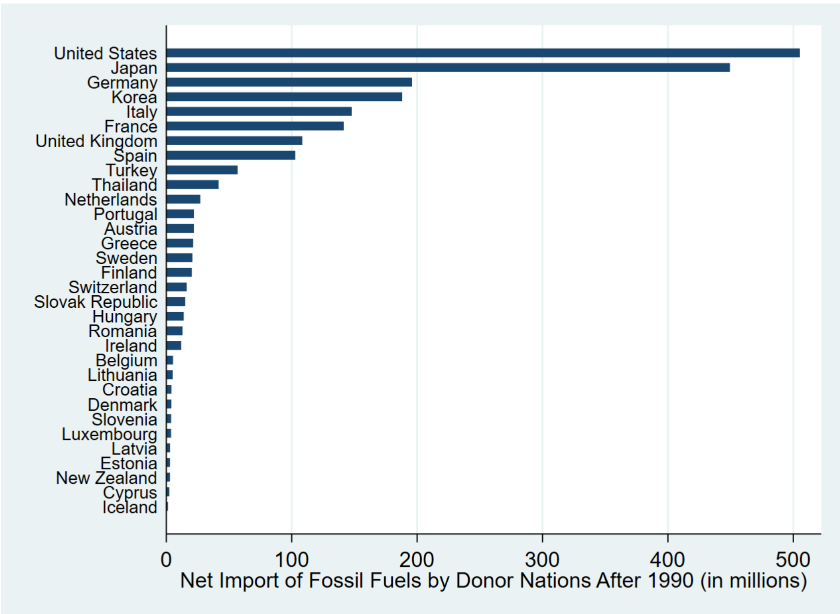
According to the formula, some resource-rich countries might have zero $log(Resources_{it})$ if their resource reserves were not yet discovered in 1990. The formula considers this group of countries to have zero value because we suspect that development aid may have promoted new resource discovery after 1990, within the period of analysis. Appendix C1.2 provides a detailed discussion of other cases of zero $log(Resources_{it})$. To further prevent endogenous bias in prices, we exclude Saudi Arabia, the top oil exporter, from the analysis.

**Demand for natural resources by donors.** We instrument the resource-access preference with the donor's demand for resource imports. The data for the instrument variable is sourced from the UNEP IRP Global Material Flows Database, which contains the physical trade balance (PTB) of natural resources by country. We compute the total net import of fossil fuels between 1991 and 2020 as a measure to proxy the donor's demand for imported natural resources.

$$
log(\text{Total fuel imports}_d) = \begin{cases} log(\sum_{t=1991}^{2020} \text{net fuel imports}_{d,t}), \\ \qquad\qquad \text{if } \sum_{t=1991}^{2020} \text{net fuel imports}_{d,t} > 0, \\ 0, \qquad otherwise. \end{cases} \tag{3.10}
$$

The choice of a time-invariant variable for the donor's demand for imported natural resources is intended to prevent any potential confounding effects from unobserved characteristics that could vary over time. Figure 3.4 illustrates the sorted measure, showing that large-economy countries are more likely to have a high preference, while Eastern European countries are at the bottom of the list. In the subsequent empirical tests, we construct two dichotomous variables to capture the donor's demand for natural resources based on the continuous measure: Top-Half Donor$d$ and Top-4 Donor$d$. These variables take the value of one if donor $d$ falls within the top half or top four of the demand distribution, respectively.

Figure 3.4: Donor's demand for natural resources




**Testing sample.** We construct a panel dataset comprising donor-recipient pairs during the period of 1990-2020. As mentioned earlier, we exclude Saudi Arabia, a major oil-producing country, from the sample to avoid potential endogenous effects that may overshadow the global market. Furthermore, we exclude observations of donors with resource abundance, as determined by zero total net fuel imports in Equation 3.10. This exclusion is to avoid aid variation due to aid-supply motivation, whereby resource-rich donors may provide more aid when the resources at home are more valuable.[11] Further-

[11]Figure 3.2 endorses this concern by showing that donor countries with more negative net imports tend to commit higher amounts of aid.

more, some donor-recipient pairs are omitted due to singleton observations, meaning they have only one year of data within the donor-recipient history or only one active donor within the recipient's record. With this setting, the total number of observations in the analysis is 60,959, covering 3,162 donor-recipient pairs during the period of 1990-2020. In some estimations that involve more control variables, the sample size decreases due to data unavailability. Table 3.1 presents the descriptive statistics of the aforementioned covariates in the sample.

Table 3.1: Descriptive Statistics

| | Mean | SD | Min | p50 | Max | N |
|---|---|---|---|---|---|---|
| DEV Aid$_{idt}$, in millions | 18.18 | 108.04 | 0.00 | 0.36 | 7,969.98 | 60,959 |
| EMR Aid$_{idt}$, in millions | 10.44 | 95.39 | 0.00 | 0.08 | 8,803.65 | 60,959 |
| $log$(Other aid$_{idt}$) | 18.61 | 1.89 | 0.00 | 18.80 | 22.85 | 60,959 |
| $log$(Resources$_{it}$) | 6.93 | 11.45 | 0.00 | 0.00 | 30.45 | 60,959 |
| $log$(Net fuel import$_d$) | 17.56 | 1.57 | 13.96 | 17.11 | 20.04 | 60,959 |
| Top-Half Donor$_d$ | 0.80 | 0.40 | 0.00 | 1.00 | 1.00 | 60,959 |
| Top-4 Donor$_d$ | 0.22 | 0.41 | 0.00 | 0.00 | 1.00 | 60,959 |
| Disaster$_{it}$ | 0.46 | 0.50 | 0.00 | 0.00 | 1.00 | 60,749 |
| $log(GDPpc_{it})$ | 7.69 | 1.03 | 5.21 | 7.75 | 11.10 | 60,959 |
| $log(GDPpc_{dt})$ | 10.52 | 0.41 | 8.70 | 10.52 | 11.63 | 60,959 |

Note: Disaster$_{it}$ is a dummy variable that equals 1 if a recipient country experiences two or more natural disasters within a year. See Appendix C1.3 for further details.

### 3.4.2 Empirical Specifications

This section outlines the empirical specifications used to test Proposition 1, examining the relationship between resource values and development aid. First, a baseline test is introduced to assess whether donors, in general, allocate more financial assistance to resource-endowed countries. The baseline test is the following.

$$Aid_{idt} = \alpha_{id} + \gamma_{dt} + \beta_0^b log(Resources_{it}) + \Gamma' X_{idt} + \epsilon_{idt}, \tag{3.11}$$

$\alpha_{id}$ and $\gamma_{dt}$ represent the donor-recipient pair and donor-time effects, respectively, $log(Resources_{it})$ denotes the resource value of recipient country $i$ at time $t$, and $\Gamma$ is a vector of coefficients corresponding to the control variables $X_{idt}$. According to the theory, the recipient's income and aid from other donors are crucial determinants of aid commitments, and are therefore included in $X_{idt}$. Donor-year effects are incorporated as stronger control instruments to facilitate the interpretation of coefficients as variations within each donor. The coefficient $\beta_0^b$, of particular interest in the baseline test, captures the overall effect of resource values on aid from donors in general.

Next, the main empirical model is presented to explore the distinction between altruistic and resource-motivated donors. We propose a fixed-effect panel regression where donor-recipient dyad fixed effects are absorbed. The empirical model is

$$\begin{aligned} Aid_{idt} = &\alpha_{id} + \gamma_{dt} + \beta_0 log(Resources_{it}) + \beta_1 log(Resources_{it})\text{Resource Demand}_d \\ &+ \Gamma' X_{idt} + \epsilon_{idt}, \end{aligned} \tag{3.12}$$

where Resource Demand$_d$ represents the donor's demand for natural resources. Controlling for fixed effects, the coefficient $\beta_1$ is the primary focus of the main test, as it captures the additional elasticity of aid commitment by resource-motivated donors to resource-rich recipients in response to changes in resource value. As outlined in Proposition 1, we expect $\beta_1$ to be positive, implying that higher resource valuations attract more aid from resource-concerned donors compared to those with lower resource demand.

To mitigate concerns regarding endogeneity bias, we construct the resource value using global market prices, thereby avoiding local determinants from either recipient or donor countries. Additionally, the inclusion of donor-year fixed effects controls for potential home-country influences that may affect dyadic aid flows. The sample selection process, as outlined in the previous subsection, further ensures that the estimations remain unaffected by major oil-exporting recipient countries or resource-rich donor nations.

In the subsequent empirical tests, we use OLS to estimate both the log-transformed aid specifications and aid percentiles across the distribution. In certain specifications, donor demand for natural resources is measured either as the continuous variable of net fossil fuel imports, or as dummy variables indicating whether a donor ranks in the top half or top four based on resource demand.

## 3.5 Empirical Results

First, we examine the baseline relationship between resource values and development aid. The regression results for Specification 3.11, along with several variations employing different outcome and control variables, are presented in Table 3.2. Overall, the findings reveal a statistically significant relationship between resource values and development aid across all specifications. Specifically, the positive coefficient of resource value suggests that higher resource values tend to attract increased development assistance. For instance, in the regression of aid percentiles displayed in Column (4), a 1% rise in resource value is associated with a 2.45 percentile increase in aid rank. This correlation corroborates the stylized fact illustrated in Figure 3.1.

Another notable result from the baseline test is the association between dyadic aid and aid from other donors. All estimations consistently show a significantly positive relationship, indicating a general trend of competitive aid, as evident in previous studies on aid fragmentation.

Table 3.2: Estimations of development aid and resource values

| | (1) Log | (2) Log | (3) Percentile | (4) Percentile |
|---|---|---|---|---|
| $log(\text{Resources}_{it})$ | 0.599*** | 0.498*** | 2.981*** | 2.453*** |
| | (0.133) | (0.130) | (0.648) | (0.628) |
| $\text{Disaster}_{it}$ | | 0.0143 | | 0.220 |
| | | (0.0488) | | (0.225) |
| $log(\text{Other aid}_{idt})$ | | 0.430*** | | 2.239*** |
| | | (0.0358) | | (0.175) |
| Constant | 6.434*** | -0.876 | 26.94*** | -11.16** |
| | (0.924) | (1.127) | (4.489) | (5.481) |
| Dyadic fixed effects | Yes | Yes | Yes | Yes |
| Donor-year fixed effects | Yes | Yes | Yes | Yes |
| Adj. R-sq | 0.643 | 0.646 | 0.708 | 0.712 |
| Donor-recipient pairs | 3162 | 3143 | 3162 | 3143 |
| Observations | 60956 | 60746 | 60956 | 60746 |

Notes: Standard errors in parentheses are clustered by donor-recipient dyad. $*p < 0.1$, $**p < 0.05$, and $***p < 0.01$. The outcome variable in Columns (1) and (2) is $\log(\text{Dev Aid}_{idt} + 1)$, while Columns (3) and (4) use the aid percentile.

Specification 3.12 examines how resource-demanding donors might respond differently from donors with less demand when the resource value changes. Table 3.3 displays the estimated coefficients of interest from the main model and several variations with different measures of demand for resources.

Across all measures of donor resource demand, the estimated coefficients consistently demonstrate that donors with higher resource demand respond more strongly to increases in resource valuation compared to donors with lower demand. For instance, the coefficient in Column (2) indicates that aid increases by 1.35 percentile points more for these high-demand donors in response to a 1% rise in resource value. In Column (3), where the dependent variable is the logarithm of aid, donors in the top half of the distribution (in terms of resource demand) commit 1.1 percentage points more aid than those in the lower half when resource values increase by 1%. Similarly, in Column (6), the top four donors increase their aid by 6 percentile points more than other donors in response to a 1% rise in resource value.

In comparison, low-demand donors, particularly those classified in the first four columns, commit less aid when resource values increase. Specifically, donors without demand (Columns 1 and 2) and those with low demand (Columns 3 and 4) allocate less aid as resource values in the destination rise. This finding suggests that an altruistic motive dominates among these low-demand donors, which aligns with the proposed theoretical framework.

Several robustness tests were conducted, with results presented in Appendix C2. The first test addresses concerns that donor demand for resources may be correlated with the donor's income. To control for this, we include the donor's GDP per capita, allowing for time variation, and interact it with the resource value in the recipient country. The results, shown in Table 3.8, indicate that the significant effects of donor resource demand remain consistent, even when accounting for the significant relationship between donor income and bilateral aid flows.

The final robustness test explores whether the higher elasticity observed in aid commitments might be driven by more severe violent conflict in recipient countries, a phenomenon often associated with the "resource curse." The primary finding of a strong direct effect persists, with no significant differences observed between countries experiencing armed conflict and those that are not. Furthermore, Appendix C2.2 contrasts the resource motivations behind development aid and emergency aid, finding no evidence of resource motivation in the latter. As expected, emergency aid is primarily mobilized in response to crises such as natural disasters or severe armed conflict, whereas development aid tends to decrease during periods of conflict.

The empirical tests primarily assess the central hypothesis concerning how resource-driven aid differs from altruistic assistance. The direct effect of resource values on development aid is both substantial and robust, even when controlling for donor income and the presence of violent conflicts in recipient countries. One interpretation of the estimates suggests that a 1% increase in a recipient country's resource value leads to a 6-percentile increase in aid commitments from the top four donors, compared to aid from other donors. This finding implies a causal relationship between development aid and resource demand, providing strong evidence of resource-access motivations among donors with high demand for natural resources.

Table 3.3: Estimations of development aid and resource values

| | (1) Log | (2) Percentile | (3) Log | (4) Percentile | (5) Log | (6) Percentile |
|---|---|---|---|---|---|---|
| $log(\text{Resources}_{it})$ | -3.662** | -21.35*** | -0.416 | -1.952 | 0.262* | 1.150* |
| | (1.590) | (7.645) | (0.300) | (1.429) | (0.140) | (0.665) |
| $log(\text{Resources}_{it}) \times log(\text{Net fuel import}_d)$ | 0.235*** | 1.347*** | | | | |
| | (0.0896) | (0.433) | | | | |
| $log(\text{Resources}_{it}) \times$ Top-Half $\text{Donor}_d$ | | | 1.099*** | 5.302*** | | |
| | | | (0.333) | (1.590) | | |
| $log(\text{Resources}_{it}) \times$ Top-4 $\text{Donor}_d$ | | | | | 1.088*** | 6.019*** |
| | | | | | (0.331) | (1.637) |
| $log(\text{Other aid}_{idt})$ | 0.430*** | 2.240*** | 0.429*** | 2.233*** | 0.431*** | 2.244*** |
| | (0.0355) | (0.173) | (0.0357) | (0.175) | (0.0356) | (0.174) |
| $\text{Disaster}_{it}$ | 0.0140 | 0.218 | 0.0144 | 0.220 | 0.0147 | 0.222 |
| | (0.0488) | (0.225) | (0.0488) | (0.225) | (0.0487) | (0.225) |
| Constant | -0.601 | -9.592* | -0.596 | -9.812* | -0.793 | -10.70** |
| | (1.121) | (5.394) | (1.122) | (5.450) | (1.108) | (5.361) |
| Dyadic fixed effects | Yes | Yes | Yes | Yes | Yes | Yes |
| Donor-year fixed effects | Yes | Yes | Yes | Yes | Yes | Yes |
| Adj. R-sq | 0.646 | 0.712 | 0.647 | 0.712 | 0.647 | 0.713 |
| Donor-recipient pairs | 3143 | 3143 | 3143 | 3143 | 3143 | 3143 |
| Observations | 60746 | 60746 | 60746 | 60746 | 60746 | 60746 |

Notes: Standard errors in parentheses are clustered by donor-recipient dyad. $*p < 0.1$, $**p < 0.05$, and $***p < 0.01$. The outcome variable in Columns (1), (3), and (5) is $\log(\text{Dev Aid}_{idt} + 1)$, while Columns (2), (4) and (6) use the aid percentile. "Top-Half Donor" refers to donors ranked in the top half of net fuel imports from 1991 to 2020. "Top-4 Donor" refers to the United States, Japan, Germany, and South Korea.

## 3.6 Estimate Donor-Specific Elasticities

In the previous empirical analysis, concerns may arise regarding the validity of donor demand for resources as an explanatory variable. Although the robustness tests attempt to mitigate this issue by controlling for donor income, other omitted factors that may be correlated with resource demand could still be influencing the results. To address these potential limitations, this section examines donor heterogeneity in response to resource values by conducting separate regression analyses.

Figure 3.5 shows the distribution of dyadic correlations between resource values and aid percentiles for each donor. Two key insights can be drawn from this figure. First, the dispersion of correlations is high even within a single donor, indicating that aid strategies devised by a donor are diverse. Second, the majority of dyadic correlations are positive, as most of the medians are above zero. This is consistent with the baseline test results in Table 3.2, which show a strong positive relationship between dyadic aid percentiles and resource values. Figure 3.6 provides examples of aid dyads with the highest correlations within the dataset.

We further calculate donor-specific elasticities toward resource values by conducting separate regressions using the baseline specification. In this simplified test, donor-year and dyadic fixed effects are replaced by year effects to avoid collinearity. The coefficients of resource values, along with their 95% confidence intervals, are shown in Figure 3.7. Most of the estimates are positive, but only a few are significantly different from zero. These countries include France, Germany, Japan, South Korea, Sweden, the United Kingdom, and the United States. Notably, except for Sweden, these countries have high resource demand, as measured by net fossil fuel imports during 1991-2020. Conversely, a few countries, such as Belgium, Italy, and Spain, exhibit a negative relationship between aid and resource values. In general, the results from the separate regressions reinforce the validity of the measure of donor demand for natural resources and effectively capture the heterogeneity among donors with respect to their resource-access motivations.

Figure 3.5: Distribution of dyadic correlations between resource values and aid percentiles

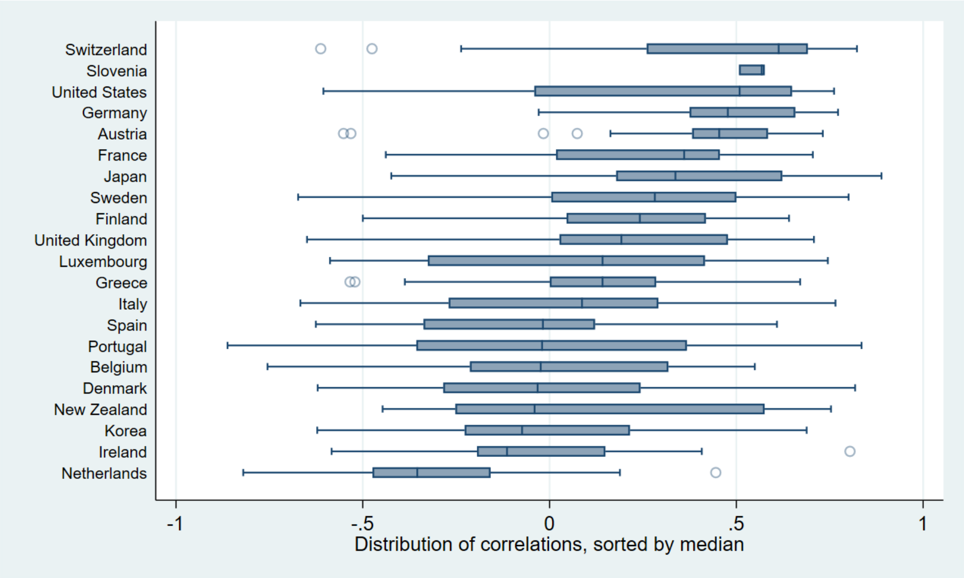



Figure 3.6: Aid contracts with the highest correlations between resource values and aid percentiles

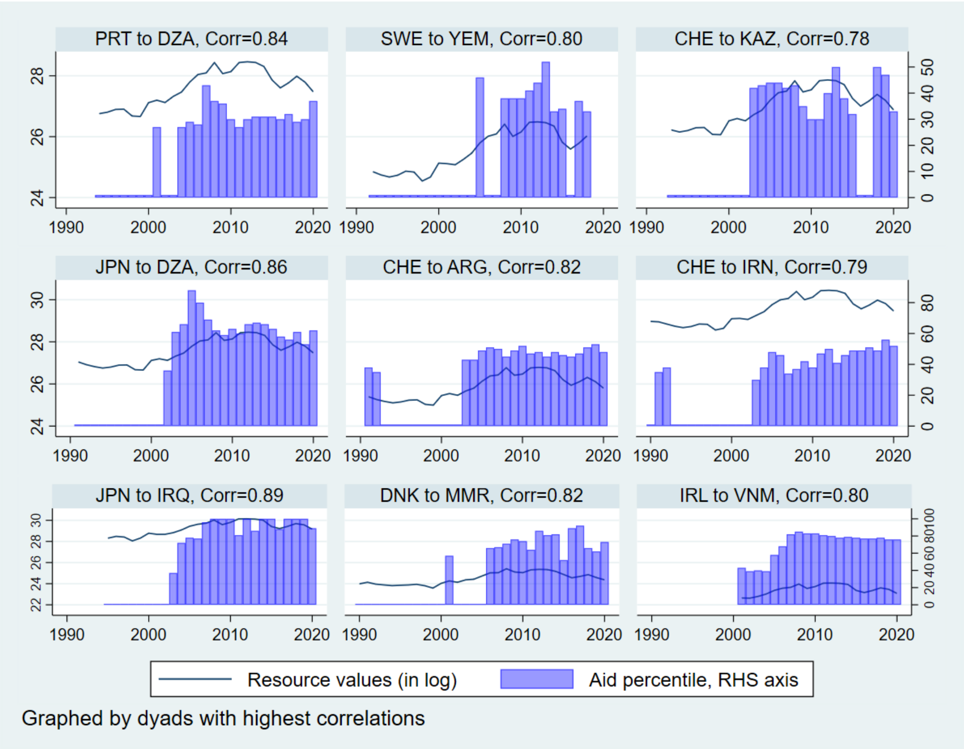
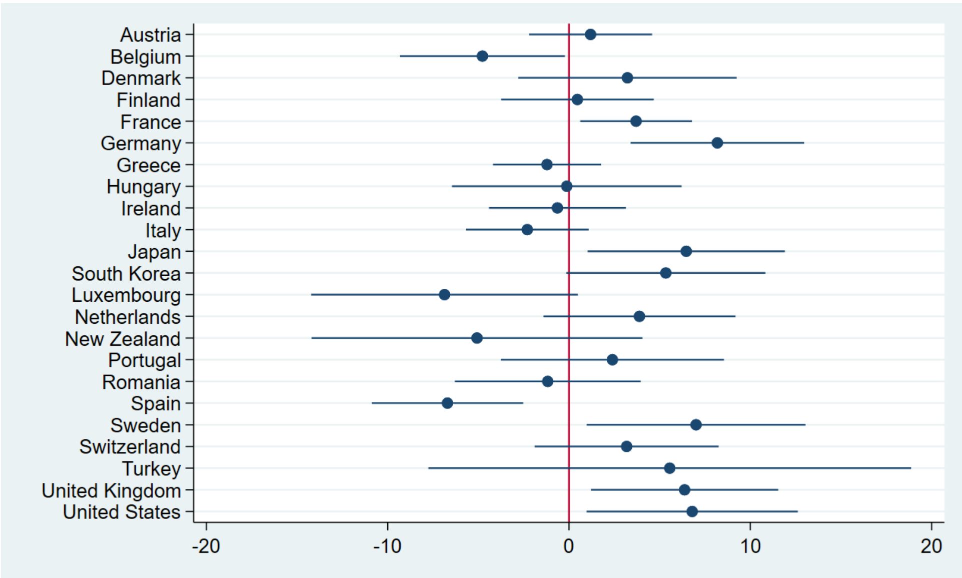


Figure 3.7: Heterogeneous aid elasticities toward resource values



## 3.7 Summary

The provision of foreign aid typically flows from higher-income to lower-income countries. However, it is notable that resource-poor nations tend to provide aid to resource-rich nations. The resource extraction industry has faced criticism, and the international community has attempted to improve its supply chain through regulations and standards. Nevertheless, we suspect that state-to-state transfers for resource access are another part of the industry that remains under-regulated and operates behind closed doors without substantive evidence to support it. Previous research has discussed instances where donor countries may provide aid in exchange for resources, but to our knowledge, only one study has explored this issue using a aid-dyad panel dataset of G7 donors, while a few others specifically study data from China and India. The literature requires further development to fully understand the issue, especially a formal framework to test the resource motivation.

We begin by examining the correlation between global foreign aid and resource prices, arguing that resource-importing donor countries tend to commit more aid. Based on qualitative studies and news reports, we believe that foreign aid systematically acts as an additional payment, on top of market compensation paid by international companies, to secure natural resources in recipient countries. We describe some fundamental concepts of all possible mechanisms through which resource values can relate to foreign aid. Given that few, if any, studies have attempted to develop a theory on this topic, we formalize this idea using an economic theory that considers both altruistic and resource-access motivations.

Our theory proposes unique predictions that identify three mechanisms of aid for resources. The first channel is the direct effect, where donor countries with higher demands for resource imports provide more development aid to support energy projects in recipient countries where their resource value valuation is higher. The second distinct signal of aid for resources is from recipient income: while development demand may cause altruistic donors to refrain from commitment, it does less to deter resource-motivated donors, as they still intend to maintain their influence in resource-rich recipients. The final channel is from aid by other donors. Similar to the income mechanism, altruistic donors coordinating development projects may refrain from providing more development aid if there are plenty of projects from other donors. However, when the resource motive is considered, resource-importing donors may respond less to increases in other aid or even intensify aid competition to protect their share of the resources. The chapter leaves the latter two hypotheses to be explore in future research.

To test our theory, we carefully construct measures to avoid endogeneity problems by relying on previous studies. Specifically, we use global fossil-fuel prices as an exogenous variation of resource values and fix the resource quantity at the time of analysis. Our novel dataset from the OECD allows us to distinguish between different types of aid by purpose and to select development aid that excludes all emergency and humanitarian aid that might be related to political violence. Empirical specifications with fixed-effect regressions are devised to absorb endogenous variation that might contaminate estimated results.

Our estimated results confirm the direct effect of resource values on dyadic aid percentiles. Specifically, the aid elasticity with respect to resource values for donors with high demand for resource imports—those in the top half of the distribution—is approximately 1.1 percentage points higher than that of other donors following a 1% increase in resource values. This causal effect remains consistent across robustness tests that control for donor income levels and account for the presence of violent conflict in the recipient country.

To address concerns regarding the measurement of donor demand, donor-specific elasticities were computed through separate regression analyses. The estimated coefficients indicate that high-demand donors, including the United States, the United Kingdom, France, Germany, Japan, and South Korea, exhibit behaviors consistent with the resource-demand measure employed in the main analysis. In contrast, only a few countries with moderate resource demand, such as Italy and Spain, display responses that diverge from theoretical expectations.

This chapter is among the earliest to explore the relationship between foreign aid motivations and resource values. Our theoretical framework offers a test to detect signals of aid motivations that contrast with altruistic intentions, and the empirical results provide quantitative evidence supporting the hypothesis of aid-for-resource motivations. A few key insights are that while the relationship between resource values and dyadic aid is generally straightforward, certain donors with moderate resource demand behave contrary to the theoretical expectations. These findings should inform future research, including further empirical tests of income and other aid-related hypotheses. Future studies may particularly focus on the efficiency of resource-driven aid in terms of global welfare and its developmental impacts.

# Appendix for Chapter 3

## C1 Additional Details on the Data

### C1.1 Aid Data from the Creditor Reporting System (CRS)

Table 3.4: Top 10 receiving countries during 1990-2020

| | DEV aid | EMR aid |
|---|---|---|
| 1. India | 93,307 | 5,828 |
| 2. China | 59,692 | 2,531 |
| 3. Indonesia | 56,686 | 23,322 |
| 4. Viet Nam | 46,988 | 6,245 |
| 5. Egypt | 42,182 | 33,724 |
| 6. Bangladesh | 39,697 | 11,161 |
| 7. Iraq | 35,940 | 64,849 |
| 8. Philippines | 35,640 | 8,042 |
| 9. Afghanistan | 34,876 | 39,875 |
| 10. Morocco | 29,685 | 3,429 |

Note: The figures are summations of aid from 1990 to 2020 in million USD at 2020 value.

Figure 3.8: Distributions of dyadic aid commitments

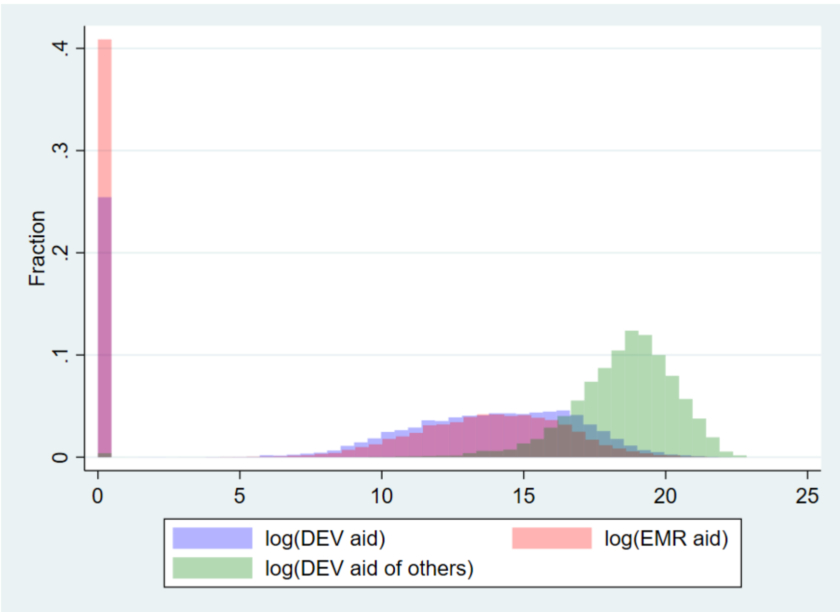
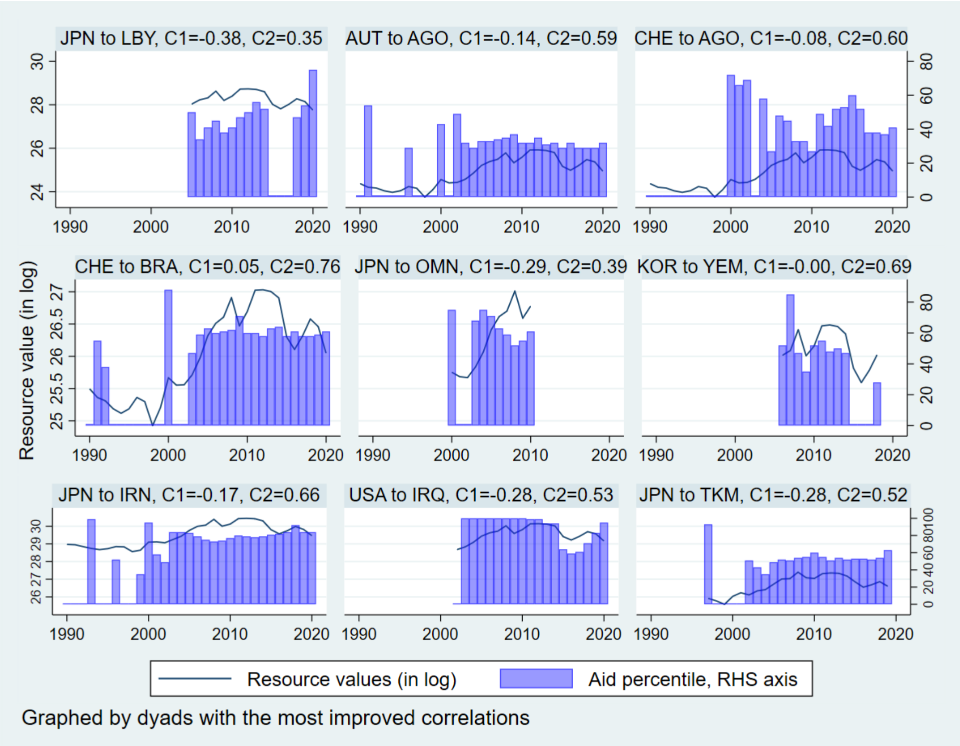


Figure 3.9: Aid contracts with the highest improved correlations from the original correlation




Note: *C1* refers to the original correlation between aid and resource values, while *C2* denotes the correlation between the aid percentiles and resource values.

## C1.2 Resource Values

The measure of resource values is constructed using three variables: a dummy variable indicating whether a country was a net fuel exporter during the period of 1991-2019, as well as the amounts of oil and gas reserves in 1990. The formula for calculating resource rents can be restated as follows:

$$log(Resources_{it}) = \mathbb{1}(\sum_{t=1991}^{2019} \text{net fossil fuel export}_{it} > 0) \times log(Q_{i,1990} * p_t),$$

where $Q_{i,1990} * p_t = (\text{Oil reserves}_{i,1990} \times Price_{oil,t}) + (\text{Gas reserves}_{i,1990} \times Price_{gas,t})$.

Table 3.5 presents a summary of the statistics of the relevant variables, including the output variable $log(Resources_{it})$, which varies across years due to changing prices. It is important to note that countries with zero $log(Resources_{it})$ can fall into one of three possible categories. The first category comprises countries with no natural reserves in 1990. The second category consists of countries that are resource producers but not enough to supply their domestic markets. Examples of such nations include China, India, and Thailand. Finally, the third category includes resource-rich countries that discovered their resource reserves after 1990. Excluding these countries from the sample of resource-rich countries prevents the potential reverse causality of development aid affecting the discovery of new resources in the recipient countries during the period of analysis. Examples of countries in this category are Mozambique, Timor-Leste, and DR Congo.

Table 3.5: Descriptive statistics of resource rents

| | Mean | SD | Min | Max | 75th PCT | N |
|---|---|---|---|---|---|---|
| $log(Resources_{it})$ | 6.27 | 11.06 | 0 | 30 | 0 | 4,428 |
| $\mathbb{1}(\sum_{t=1991}^{2019} \text{net fossil fuel export}_{it} > 0)$ | 0.33 | 0.47 | 0 | 1 | 1 | 165 |
| Oil reserves$_{i,1990}$ | 3.60 | 15.85 | 0 | 100 | 0 | 165 |
| Gas reserves$_{i,1990}$ | 12,065.42 | 52,187.48 | 0 | 583,901 | 0 | 165 |

Notes: Units of Oil reserves$_{i,1990}$ and Gas reserves$_{i,1990}$ are billion barrels and Trillion BTU. Statistics of $log(Resources_{it})$ are computed from the cross-country panel data from 1990-2020 while the sample size of $\mathbb{1}(\sum_{t=1991}^{2019} \text{net fossil fuel export}_{it} > 0)$, Oil reserves$_{i,1990}$ and Gas reserves$_{i,1990}$ is the number of receiving countries.

Table 3.6: Countries with the highest oil reserves in 1990

| | Oil reserves (in billion barrels) |
|---|---|
| 1. Iraq | 100.00 |
| 2. United Arab Emirates | 98.10 |
| 3. Kuwait | 97.03 |
| 4. Iran | 92.85 |
| 5. Mexico | 26.90 |
| 6. Libya | 22.80 |
| 7. Nigeria | 17.10 |
| 8. Algeria | 9.20 |
| 9. Indonesia | 5.42 |
| 10. Kazakhstan | 5.19 |

Note: Only countries included in the sample.

Table 3.7: Countries with the highest gas reserves in 1990

| | Gas reserves (in TBtu) |
|---|---|
| 1. Iran | 583,901.3 |
| 2. United Arab Emirates | 199,335.5 |
| 3. Algeria | 115,485.3 |
| 4. Iraq | 107,319.0 |
| 5. Indonesia | 105,679.4 |
| 6. Nigeria | 98,096.6 |
| 7. Mexico | 71,786.3 |
| 8. Turkmenistan | 63,916.4 |
| 9. Kazakhstan | 62,291.4 |
| 10. Kuwait | 52,433.3 |

Note: Only countries included in the sample.

## C1.3 Natural Disasters as Exogenous Income Shocks

**Recipient's exogenous income shock.** In accordance with the common usage in the literature, we measure the exogenous variation in income by incorporating a natural disaster variable obtained from the Emergency Events Database (EM-DAT) (Besley and Persson, 2011). The concept of natural disaster captures the adverse impact on a country's national income and population welfare resulting from natural or technological events, which are defined as any of the following:

1. 10 or more fatalities;
2. 100 or more people affected;
3. The declaration of a state of emergency;
4. A call for international assistance.

In this study, with a slight modification, we construct a binary variable $\text{Disaster}_{it}$, which takes the value of 1 if a recipient country experiences two or more natural disasters in a given year, and 0 otherwise. This measure reflects an exogenous negative income shock to the recipient country. The summary statistics in Table 3.1 indicate that 46 percent of the dyadic cases involve at least two natural disasters.

# C2 Robustness Tests

## C2.1 Foreign aid and donor's income

Table 3.8: Estimations of aid percentiles, resource values, and donor's income

| | (1) Log | (2) Percentile | (3) Log | (4) Percentile | (5) Log | (6) Percentile |
|---|---|---|---|---|---|---|
| $log(\text{Resources}_{it})$ | -4.689*** | -26.70*** | -1.095** | -5.486*** | -0.581 | -3.224* |
| | (1.660) | (8.000) | (0.443) | (2.097) | (0.366) | (1.737) |
| $log(\text{Resources}_{it}) \times log(\text{Net fuel import}_d)$ | 0.245*** | 1.395*** | | | | |
| | (0.0899) | (0.436) | | | | |
| $log(\text{Resources}_{it}) \times$ Top-Half $\text{Donor}_d$ | | | 1.070*** | 5.149*** | | |
| | | | (0.333) | (1.593) | | |
| $log(\text{Resources}_{it}) \times$ Top-4 $\text{Donor}_d$ | | | | | 1.107*** | 6.117*** |
| | | | | | (0.333) | (1.646) |
| $log(\text{Resources}_{it}) \times log(GDPpc_{dt})$ | 0.0648** | 0.338** | 0.0529* | 0.275* | 0.0631** | 0.327** |
| | (0.0296) | (0.143) | (0.0295) | (0.143) | (0.0295) | (0.143) |
| $log(\text{Other aid}_{idt})$ | 0.428*** | 2.225*** | 0.427*** | 2.221*** | 0.428*** | 2.230*** |
| | (0.0355) | (0.174) | (0.0358) | (0.175) | (0.0356) | (0.174) |
| $\text{Disaster}_{it}$ | 0.0119 | 0.207 | 0.0127 | 0.211 | 0.0126 | 0.211 |
| | (0.0487) | (0.225) | (0.0487) | (0.225) | (0.0486) | (0.225) |
| Constant | 0.723 | -2.688 | 0.468 | -4.277 | 0.486 | -4.069 |
| | (1.077) | (5.092) | (1.080) | (5.173) | (1.056) | (5.031) |
| Dyadic fixed effects | Yes | Yes | Yes | Yes | Yes | Yes |
| Donor-year fixed effects | Yes | Yes | Yes | Yes | Yes | Yes |
| Adj. R-sq | 0.647 | 0.713 | 0.647 | 0.712 | 0.647 | 0.713 |
| Donor-recipient pairs | 3143 | 3143 | 3143 | 3143 | 3143 | 3143 |
| Observations | 60746 | 60746 | 60746 | 60746 | 60746 | 60746 |

Notes: Standard errors in parentheses are clustered by donor-recipient dyad. $*p < 0.1$, $**p < 0.05$, and $***p < 0.01$. The outcome variable in Columns (1), (3), and (5) is $\log(\text{Dev Aid}_{idt} + 1)$, while Columns (2), (4) and (6) use the aid percentile. "Top-Half Donor" refers to donors ranked in the top half of net fuel imports from 1991 to 2020. "Top-4 Donor" refers to the United States, Japan, Germany, and South Korea.

## C2.2 Foreign aid and political violence

This appendix section presents a comparison of the correlation between development aid (*DEV Aid*) and emergency aid (*EMR Aid*) with political violence. The political violence data is obtained from the Georeferenced Event Dataset (GED) Global version 20.1 from Uppsala Conflict Data Program (UCDP), which is widely used to study cross-country conflict. The GED dataset has broad coverage of countries and time periods, including conflict events that cross the threshold of 25 deaths in any given year (Sundberg and Melander, 2013). In this study, the variable *best* is used, which represents the best estimate of total fatalities resulting from conflict events. The measure employed for the estimation is $log(fatalities_{it}) \equiv log(best_{it} + 1)$.

Table 3.9 presents Poisson regressions that indicate different relationships between development and emergency aid on armed conflicts. Columns (1) and (2) show that development aid diminishes when a recipient country is experiencing severe armed conflict, while emergency aid acts as a responsive assistance to the conflict. Additionally, the results from these first two columns also serve as a robustness test for the main findings by including armed conflicts in the estimation. The coefficient of the resource rent for high-resource-demanding donors remains significant and positive.

Table 3.9: Estimations of aid percentiles, resource values, and armed conflicts

| | (1)<br>DEV Aid | (2)<br>DEV Aid | (3)<br>EMR Aid | (4)<br>EMR aid |
|---|---|---|---|---|
| $log(\text{Resources}_{it})$ | -0.0367 | 0.0313* | 0.0886* | 0.0843*** |
| | (0.0378) | (0.0162) | (0.0538) | (0.0208) |
| $log(\text{Other aid}_{idt})$ | 0.0671*** | 0.0672*** | 0.0402*** | 0.0398*** |
| | (0.00509) | (0.00508) | (0.00584) | (0.00585) |
| i.noDisaster$_{it}$ | -0.00151 | -0.00160 | -0.0345*** | -0.0344*** |
| | (0.00526) | (0.00526) | (0.00741) | (0.00741) |
| $log(\text{Fatalities}_{it})$ | -0.00537*** | -0.00533*** | 0.0207*** | 0.0207*** |
| | (0.00205) | (0.00205) | (0.00241) | (0.00241) |
| $log(\text{Resources}_{it}) \times log(\text{Fatalities}_{it})$ | 0.000293 | 0.000482*** | 0.00212*** | 0.00168*** |
| | (0.000369) | (0.000162) | (0.000608) | (0.000258) |
| $log(\text{Resources}_{it}) \times$ i.Top-Half Donor$_d$ | 0.103** | | -0.00288 | |
| | (0.0407) | | (0.0569) | |
| $log(\text{Resources}_{it}) \times$ i.Top-Half Donor$_d \times log(\text{Fatalities}_{it})$ | 0.000115 | | -0.000946 | |
| | (0.000386) | | (0.000637) | |
| $log(\text{Resources}_{it}) \times$ i.Top-4 Donor$_d$ | | 0.0761** | | 0.00576 |
| | | (0.0312) | | (0.0382) |
| $log(\text{Resources}_{it}) \times$ i.Top-4 Donor$_d \times log(\text{Fatalities}_{it})$ | | -0.000275 | | -0.00128*** |
| | | (0.000298) | | (0.000403) |
| Constant | 2.429*** | 2.415*** | 2.575*** | 2.585*** |
| | (0.139) | (0.138) | (0.172) | (0.170) |
| Dyadic fixed effects | Yes | Yes | Yes | Yes |
| Donor-year fixed effects | Yes | Yes | Yes | Yes |
| Pseudo-R-sq | 0.534 | 0.535 | 0.500 | 0.500 |
| Donor-recipient pairs | 3100 | 3100 | 3100 | 3100 |
| Observations | 58353 | 58353 | 58353 | 58353 |

Notes: Standard errors in parentheses are clustered by donor-recipient dyad. $*p < 0.1$, $**p < 0.05$, and $***p < 0.01$. The dependent variable is the aid percentile, with Columns (1) and (2) representing development aid, and Columns (3) and (4) representing emergency aid. "Top-Half Donor" refers to donors ranked in the top half of net fuel imports from 1991 to 2020, while "Top-4 Donor" denotes the United States, Japan, Germany, and South Korea.

# Bibliography


Abadie, A. and Gardeazabal, J. (2003). The economic costs of conflict: A case study of the basque country. *American Economic Review*, 93(1):113–132.

ADB (2020). *Mapping Poverty through Data Integration and Artificial Intelligence: A Special Supplement of the Key Indicators for Asia and the Pacific*. Asian Development Bank.

Ahmed, F. Z. (2012). The perils of unearned foreign income: Aid, remittances, and government survival. *The American Political Science Review*, 106(1):146–165.

Ahmed, F. Z. and Werker, E. D. (2015). Aid and the rise and fall of conflict in the muslim world. *Quarterly Journal of Political Science*, 10(2):155–186.

Akbulut-Yuksel, M. (2014). Children of war: the long-run effects of large-scale physical destruction and warfare on children. *Journal of Human Resources*, 49(3):634–662.

Akresh, R., Bhalotra, S., Leone, M., and Osili, U. O. (2012). War and stature: Growing up during the nigerian civil war. *American Economic Review*, 102(3):273–277.

Alesina, A. and Dollar, D. (2000). Who gives foreign aid to whom and why? *Journal of Economic Growth*, 5(1):33–63.

Alkire, S. and Jahan, S. (2018). *The New Global MPI 2018: Aligning with the Sustainable Development Goals*. UNDP.

Amodio, F., Baccini, L., and Di Maio, M. (2020). Security, trade, and political violence. *Journal of the European Economic Association*.

Andersen, J. J., Nordvik, F. M., and Tesei, A. (2022). Oil price shocks and conflict escalation: Onshore versus offshore. *Journal of Conflict Resolution*, 66(2):327–356.

Annen, K. and Moers, L. (2017). Donor competition for aid impact, and aid fragmentation. *World Bank Economic Review*, 31(3):708–729.

Asatullaeva, Z., Aghdam, R. F. Z., Ahmad, N., and Tashpulatova, L. (2021). The impact of foreign aid on economic development: A systematic literature review and content analysis of the top 50 most influential papers. *Journal of International Development*, 33(4):717–751.

Babenko, B., Hersh, J., Newhouse, D., Ramakrishnan, A., and Swartz, T. (2017). Poverty mapping using convolutional neural networks trained on high and medium resolution satellite images, with an application in mexico. *arXiv preprint arXiv:1711.06323*.

Barro, R. J. and Sala-i Martin, X. I. (2003). *Economic Growth, Second Edition*. The MIT Press.

Bazzi, S., Blair, R. A., Blattman, C., Dube, O., Gudgeon, M., and Peck, R. M. (2019). The promise and pitfalls of conflict prediction: Evidence from colombia and indonesia. Technical report, National Bureau of Economic Research.

Behrman, J. R. and Rosenzweig, M. R. (2004). Returns to birthweight. *Review of Economics and Statistics*, 86(2):586–601.

Berman, N., Couttenier, M., Rohner, D., and Thoenig, M. (2017). This mine is mine! how minerals fuel conflicts in africa. *American Economic Review*, 107(6):1564–1610.

Besley, T., Fetzer, T., and Mueller, H. (2020). Terror and tourism: The economic consequences of media coverage. Technical report, Barcelona GSE Working Paper 1141.

Besley, T. and Mueller, H. (2012). Estimating the peace dividend: The impact of violence on house prices in northern ireland. *American Economic Review*, 102(2):810–833.

Besley, T. and Mueller, H. (2018). Predation, protection, and productivity: A firm-level perspective. *American Economic Journal: Macroeconomics*, 10(2):184–221.

Besley, T. and Mueller, H. (2020). *Revenue Mobilization and State Fragility*, pages 235–270. Oxford University Press.

Besley, T. and Persson, T. (2011). The logic of political violence. *The Quarterly Journal of Economics*, 126(3):1411–1445.

Beyer, R. C., Hu, Y., and Yao, J. (2022). *Measuring quarterly economic growth from outer space*. Number 9893. International Monetary Fund.

Blattman, C. and Annan, J. (2010). The consequences of child soldiering. *The Review of Economics and Statistics*, 92(4):882–898.

Bok, B., Caratelli, D., Giannone, D., Sbordone, A. M., and Tambalotti, A. (2018). Macroeconomic nowcasting and forecasting with big data. *Annual Review of Economics*, 10(1):615–643.

Bormann, N.-C., Cederman, L.-E., Gates, S., Graham, B. A. T., Hug, S., Strøm, K. W., and Wucherpfennig, J. (2019). Power sharing: Institutions, behavior, and peace. *American Journal of Political Science*, 63(1):84–100.

BP (2021). Bp statistical review of world energy. Technical report, BP.

Brakman, S., Garretsen, H., and Schramm, M. (2004). The strategic bombing of german cities during world war ii and its impact on city growth. *Journal of Economic Geography*, 4(2):201–218.

Brand, S. and Price, R. (2000). *The Economic and Social Costs of Crime*. Home Office.

Brazys, S., Elkink, J. A., and Kelly, G. (2017). Bad neighbors? How co-located Chinese and World Bank development projects impact local corruption in Tanzania. *The Review of International Organizations*, 12(2):227–253.

Buell, B., Cherif, R., Chen, C., Walentin, K., Tang, J., and Wendt, N. (2021). *Impact of COVID-19: Nowcasting and big data to track economic activity in Sub-Saharan Africa*. International Monetary Fund.

Burnside, C. and Dollar, D. (2000). Aid, policies, and growth. *American Economic Review*, 90(4):847–868.

Cameron, G. J., Dang, H.-A. H., Dinc, M., Foster, J. S., and Lokshin, M. M. (2019). Measuring the Statistical Capacity of Nations. Policy Research Working Paper Series 8693, The World Bank.

Case, A. and Paxson, C. (2008). Stature and status: Height, ability, and labor market outcomes. *Journal of Political Economy*, 116(3):499–532.

Cerra, V. and Saxena, S. C. (2008). Growth dynamics: The myth of economic recovery. *American Economic Review*, 98(1):439–457.

Chen, J. and Roth, J. (2023). Logs with Zeros? Some Problems and Solutions*. *The Quarterly Journal of Economics*, 139(2):891–936.

Chen, W., Chen, X., Hsieh, C.-T., and Song, Z. (2019). A forensic examination of china's national accounts. *Brookings Papers on Economic Activity*, pages 77–127.

Chen, X. and Nordhaus, W. D. (2011). Using luminosity data as a proxy for economic statistics. *Proceedings of the National Academy of Sciences*, 108(21):8589–8594.

Choi, H. and Varian, H. (2012). Predicting the present with google trends. *Economic Record*, 88(s1):2–9.

Collier, P. (1999). On the economic consequences of civil war. *Oxford Economic Papers*, 51(1):168–183.

Collier, P. and Dehn, J. (2001). Aid, shocks, and growth. Policy Research Working Paper Series 2688, The World Bank.

Collier, P. and Dollar, D. (2002). Aid allocation and poverty reduction. *European Economic Review*, 46(8):1475–1500.

Collier, P. and Hoeffler, A. (2004). Greed and grievance in civil war. *Oxford Economic Papers*, 56(4):563–595.

Commission, E. and Centre, J. R. (2023). *GHSL data package 2023*. Publications Office of the European Union.

Cook, P. J. and MacDonald, J. (2011). Public safety through private action: An economic assessment of bids. *Economic Journal*, 121(552):445–462.

Corral, P., Irwin, A., Krishnan, N., and Mahler, D. G. (2020). *Fragility and Conflict: On the Front Lines of the Fight Against Poverty*. World Bank, Washington, DC.

Couharde, C., Karanfil, F., Kilama, E. G., and Omgba, L. D. (2020). The role of oil in the allocation of foreign aid: The case of the g7 donors. *Journal of Comparative Economics*, 48(2):363–383.

Davis, D. R. and Weinstein, D. E. (2002). Bones, bombs, and break points: The geography of economic activity. *American Economic Review*, 92(5):1269–1289.

de Mesquita, B. B. and Smith, A. (2009). A political economy of aid. *International Organization*, 63(2):309–340.

Donaldson, D. and Storeygard, A. (2016). The view from above: Applications of satellite data in economics. *Journal of Economic Perspectives*, 30(4):171–98.

Dreher, A., Eichenauer, V. Z., and Gehring, K. (2016). Geopolitics, Aid, and Growth: The Impact of UN Security Council Membership on the Effectiveness of Aid. *The World Bank Economic Review*, 32(2):268–286.

Dreher, A. and Fuchs, A. (2015). Rogue aid? an empirical analysis of china's aid allocation. *Canadian Journal of Economics/Revue canadienne d'économique*, 48(3):988–1023.

Dreher, A., Fuchs, A., Parks, B., Strange, A., and Tierney, M. J. (2021). Aid, china, and growth: Evidence from a new global development finance dataset. *American Economic Journal: Economic Policy*, 13(2):135–74.

Dreher, A., Fuchs, A., Parks, B., Strange, A. M., and Tierney, M. J. (2018). Apples and Dragon Fruits: The Determinants of Aid and Other Forms of State Financing from China to Africa. *International Studies Quarterly*, 62(1):182–194.

Dreher, A. and Lohmann, S. (2015). Aid and growth at the regional level. *Oxford Review of Economic Policy*, 31(3-4):420–446.

Dreher, A., Nunnenkamp, P., and Thiele, R. (2011). Are 'new' donors different? comparing the allocation of bilateral aid between nondac and dac donor countries. *World Development*, 39(11):1950–1968. Expanding Our Understanding of Aid with a New Generation in Development Finance Information.

Dreher, A., Sturm, J.-E., and Vreeland, J. R. (2009). Development aid and international politics: Does membership on the un security council influence world bank decisions? *Journal of Development Economics*, 88(1):1–18.

Dube, O. and Vargas, J. F. (2013). Commodity Price Shocks and Civil Conflict: Evidence from Colombia. *The Review of Economic Studies*, 80(4):1384–1421.

Duflo, E. (2001). Schooling and labor market consequences of school construction in indonesia: Evidence from an unusual policy experiment. *American Economic Review*, 91(4):795–813.

Easterly, W. (2003). Can foreign aid buy growth? *Journal of Economic Perspectives*, 17(3):23–48.

Economist, T. (2015). Fear itself. *The Economist*. Accessed May 15, 2020.

EIA (2019). International energy outlook 2019. Technical report, U.S. Energy Information Administration.

Elvidge, C. D., Baugh, K., Zhizhin, M., Hsu, F. C., and Ghosh, T. (2017). Viirs nighttime lights. *International Journal of Remote Sensing*, 38(21):5860–5879.

Fetzer, T. and Kyburz, S. (2018). Cohesive institutions and political violence. Technical report, CAGE Online Working Paper Series 377, Competitive Advantage in the Global Economy (CAGE).

Friedman, M. J. and Schnurr, P. P. (1995). *The Relationship between Trauma, Post-Traumatic Stress Disorder, and Physical Health*, pages 507–524. Lippincott Williams & Wilkins, Philadelphia.

Fuchs, A. and Vadlamannati, K. C. (2013). The needy donor: An empirical analysis of india's aid motives. *World Development*, 44:110 – 128.

Galiani, S., Knack, S., Xu, L., and Zou, B. (2017). The effect of aid on growth: evidence from a quasi-experiment. *Journal of Economic Growth*, 22(1):1–33.

Gehring, K., Kaplan, L. C., and Wong, M. H. (2019). China and the world bank: How contrasting development approaches affect the stability of african states. *CESifo Working Paper*.

Gehring, K., Michaelowa, K., Dreher, A., and Spörri, F. (2017). Aid fragmentation and effectiveness: What do we really know? *World Development*, 99:320–334.

Gibbons, S. (2004). The costs of urban property crime. *Economic Journal*, 114(499):F441–F463.

Gibson, J., Olivia, S., and Boe-Gibson, G. (2020). Night lights in economics: Sources and uses1. *Journal of Economic Surveys*, 34(5):955–980.

Hackenesch, C. (2015). Not as bad as it seems: Eu and us democracy promotion faces china in africa. *Democratization*, 22(3):419–437.

Hegre, H., Allansson, M., Basedau, M., Colaresi, M., Croicu, M., Fjelde, H., and Hoyles, F. e. a. (2019). Views: A political violence early-warning system. *Journal of Peace Research*, 56(2):155–174.

Henderson, J. V., Squires, T., Storeygard, A., and Weil, D. (2017). The Global Distribution of Economic Activity: Nature, History, and the Role of Trade1. *The Quarterly Journal of Economics*, 133(1):357–406.

Henderson, J. V., Storeygard, A., and Weil, D. N. (2012). Measuring economic growth from outer space. *American Economic Review*, 102(2):994–1028.

Hernandez, D. (2017). Are "new" donors challenging world bank conditionality? *World Development*, 96(C):529–549.

Hernandez, M., Hong, L., Frias-Martinez, V., Whitby, A., and Frias-Martinez, E. (2017). Estimating poverty using cell phone data: evidence from guatemala. *World Bank Policy Research Working Paper*, (7969).

Herrendorf, B., Rogerson, R., and Valentinyi, A. (2014). Growth and structural transformation. *Handbook of economic growth*, 2:855–941.

Hodler, R. and Raschky, P. A. (2014). Regional Favoritism *. *The Quarterly Journal of Economics*, 129(2):995–1033.

Hoeffler, A. and Fearon, J. (2015). *Post-2015 Consensus: Conflict and Violence Assessment*. Copenhagen Consensus Center, Copenhagen.

Hoeffler, A. and Outram, V. (2011). Need, merit, or self-interest—what determines the allocation of aid? *Review of Development Economics*, 15(2):237–250.

Hühne, P., Meyer, B., and Nunnenkamp, P. (2014). Who benefits from aid for trade? comparing the effects on recipient versus donor exports. *The Journal of Development Studies*, 50(9):1275–1288.

Ichino, A. and Winter-Ebmer, R. (2004). The long-run educational cost of world war ii. *Journal of Labor Economics*, 22(1):57–86.

Jean, N., Burke, M., Xie, M., Davis, W. M., Lobell, D. B., and Ermon, S. (2016). Combining satellite imagery and machine learning to predict poverty. *Science*, 353(6301):790–794.

Kaya, I. and Kaya, O. (2020). Foreign aid, institutional quality and government fiscal behavior in emerging economies: An empirical investigation. *The Quarterly Review of Economics and Finance*, 76:59–67.

Khachiyan, A., Thomas, A., Zhou, H., Hanson, G., Cloninger, A., Rosing, T., and Khandelwal, A. K. (2022). Using neural networks to predict microspatial economic growth. *American Economic Review: Insights*, 4(4):491–506.

Kummu, M., Taka, M., and Guillaume, J. H. (2018). Gridded global datasets for gross domestic product and human development index over 1990–2015. *Scientific data*, 5(1):1–15.

Kuziemko, I. and Werker, E. (2006). How much is a seat on the security council worth? foreign aid and bribery at the united nations. *Journal of Political Economy*, 114(5):905–930.

León, G. (2012). Civil conflict and human capital accumulation: The long-term effects of political violence in peru. *Journal of Human Resources*, 47(4):991–1022.

Li, B., Mahan, C. M., Kang, H. K., Eisen, S. A., and Engel, C. C. (2011). Longitudinal health study of us 1991 gulf war veterans: Changes in health status at 10-year follow-up. *American Journal of Epidemiology*, 174(7):761–768.

Li, X. (2017). Does conditionality still work? china's development assistance and democracy in africa. *Chinese Political Science Review*, 2(2):201–220.

Lum, T., Fischer, H., Gomez-Granger, J., and Leland, A. (2009). China's foreign aid activities in africa, latin america, and southeast asia. *Russia China and Eurasia-Social Historicaland Cultural Issues*, 25(2):175.

Maruta, A. A., Banerjee, R., and Cavoli, T. (2020). Foreign aid, institutional quality and economic growth: Evidence from the developing world. *Economic Modelling*, 89:444–463.

Marx, A. and Soares, J. (2013). South korea's transition from recipient to dac donor: Assessing korea's development cooperation policy. *International Development Policy—Revue internationale de politique de développement*, pages 107–142.

Mathur, J. (2020). predicting-poverty-replication.

Miguel, E. and Roland, G. (2011). The long-run impact of bombing vietnam. *Journal of Development Economics*, 96(1):1–15.

Minoiu, C. and Shemyakina, O. N. (2014). Armed conflict, household victimization, and child health in côte d'ivoire. *Journal of Development Economics*, 108:237–255.

Mueller, H. (2012). Growth dynamics: The myth of economic recovery: Comment. *American Economic Review*, 102(7):3774–3777.

Mueller, H. (2016). Growth and violence: Argument for a per capita measure of civil war. *Economica*, 83(331):473–497.

Mueller, H., Piemontese, L., and Tapsoba, A. (2017). *Recovery from Conflict: Lessons of Success*. World Bank, Washington.

Mueller, H. and Rauh, C. (2019). The hard problem of prediction for conflict prevention. Technical report, CEPR Discussion Papers 13748.

Mueller, H. and Rohner, D. (2018). Can power-sharing foster peace? evidence from northern ireland. *Economic Policy*, 33(95):447–484.

Mueller, H. and Tobias, J. (2016). The cost of violence: Estimating the economic impact of conflict. Technical report, International Growth Centre (IGC).

Narita, F. and Yin, R. (2018). In search of information: Use of google trends' data to narrow information gaps for low-income developing countries*. *IMF Working Papers*, 2018(286):A001.

Nordhaus, W. and Chen, X. (2016). Global gridded geographically based economic data (g-econ), version 4. *NASA Socioeconomic Data and Applications Center (SEDAC).*

OECD (2022). Net oda (indicator).

Organski, A. F. K. and Kugler, J. (1977). The costs of major wars: The phoenix factor. *American Political Science Review*, 71(4):1347–1366.

Pettersson, J. and Johansson, L. (2013). Aid, aid for trade, and bilateral trade: An empirical study. *The Journal of International Trade & Economic Development*, 22(6):866–894.

Price, N. and Atkinson, P. M. (2022). Global gdp prediction with night-lights and transfer learning. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 15:7128–7138.

Puttanapong, N., Martinez, A., Bulan, J. A. N., Addawe, M., Durante, R. L., and Martillan, M. (2022). Predicting poverty using geospatial data in thailand. *ISPRS International Journal of Geo-Information*, 11(5).

Rajan, R. G. and Subramanian, A. (2008). Aid and growth: What does the cross-country evidence really show? *The Review of Economics and Statistics*, 90(4):643–665.

Reilly, J. (2013). China and japan in myanmar: Aid, natural resources and influence. *Asian Studies Review*, 37(2):141–157.

Rolf, E., Proctor, J., Carleton, T., Bolliger, I., Shankar, V., Ishihara, M., Recht, B., and Hsiang, S. (2021). A generalizable and accessible approach to machine learning with global satellite imagery. *Nature communications*, 12(1):4392.

Rosen, J. and Fields, R. (1988). The long-term effects of extraordinary trauma: A look beyond ptsd. *Journal of Anxiety Disorders*, 2(2):179–191.

Sanchez de la Sierra, R. (2020). On the origins of the state: Stationary bandits and taxation in eastern congo. *Journal of Political Economy*, pages 32–74.

Schiavina, M., Melchiorri, M., and Pesaresi, M. (2023). Ghs-smod r2023a - ghs settlement layers, application of the degree of urbanisation methodology (stage i) to ghs-pop r2023a and ghs-built-s r2023a, multitemporal (1975-2030).

Sibai, A. M., Armenian, H. K., and Alam, S. (1989). Wartime determinants of arteriographically confirmed coronary artery disease in beirut. *American Journal of Epidemiology*, 130(4):623–631.

Singhal, S. (2019). Early life shocks and mental health: The long-term effect of war in vietnam. *Journal of Development Economics*, 141:102244.

Soares, R. (2009). *Welfare Costs of Crime and Common Violence: A Critical Review*, pages 27–47. World Bank, Washington, DC.

Steinwand, M. C. (2015). Compete or coordinate? aid fragmentation and lead donorship. *International Organization*, 69(2):443–472.

Strauss, J. and Thomas, D. (1998). Health, nutrition, and economic development. *Journal of Economic Literature*, 36(2):766–817.

Sundberg, R. and Melander, E. (2013). Introducing the ucdp georeferenced event dataset. *Journal of Peace Research*, 50(4):523–532.

Sutton, P. C., Elvidge, C. D., Ghosh, T., et al. (2007). Estimation of gross domestic product at sub-national scales using nighttime satellite imagery. *International Journal of Ecological Economics & Statistics*, 8(S07):5–21.

Tapsoba, A. (2018). The cost of fear: Impact of violence risk on child health during conflict. Technical report, HiCN Working Papers No 279, Households in Conflict Network.

The LSE-Oxford Commission on State Fragility, G. and Development (2018). *Escaping the Fragility Trap*. International Growth Centre (IGC), London.

Tingzon, I., Orden, A., Go, K. T., Sy, S., Sekara, V., Weber, I., Fatehkia, M., García-Herranz, M., and Kim, D. (2019). Mapping poverty in the philippines using machine learning, satellite imagery, and crowd-sourced geospatial information. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLII-4/W19:425–431.

Tullock, G. (2001). Efficient rent seeking. In Lockard, A. A. and Tullock, G., editors, *Efficient Rent-Seeking: Chronicle of an Intellectual Quagmire*, pages 3–16. Springer US.

van Ours, J. C. and Vollaard, B. (2016). The engine immobiliser: A non-starter for car thieves. *Economic Journal*, 126(593):1264–1291.

Wang, X., Sutton, P. C., and Qi, B. (2019). Global mapping of gdp at 1 km2 using viirs nighttime satellite imagery. *ISPRS International Journal of Geo-Information*, 8(12):580.

Watt, J. (2019). £680m of uk foreign aid spent on fossil fuel projects – study.

Wenz, L., Carr, R. D., Kögel, N., Kotz, M., and Kalkuhl, M. (2023). Dose–global data set of reported sub-national economic output. *Scientific Data*, 10(1):425.

Werker, E., Ahmed, F. Z., and Cohen, C. (2009). How is foreign aid spent? evidence from a natural experiment. *American Economic Journal: Macroeconomics*, 1(2):225–44.

Willard, K. L., Guinnane, T. W., and Rosen, H. S. (1996). Turning points in the civil war: Views from the greenback market. *American Economic Review*, 86(4):1001–1018.

Wolfe, J., Schnurr, P. P., Brown, P. J., and Furey, J. (1994). Posttraumatic stress disorder and war-zone exposure as correlates of perceived health in female vietnam war veterans. *Journal of Consulting and Clinical Psychology*, 62(6):1235–1240.

World Bank (2017). *The Toll of War: The Economic and Social Consequences of the Conflict in Syria*. World Bank, Washington, DC.

Yeh, C., Perez, A., Driscoll, A., Azzari, G., Tang, Z., Lobell, D., Ermon, S., and Burke, M. (2020). Using publicly available satellite imagery and deep learning to understand economic well-being in africa. *Nature communications*, 11(1):2583.

Zussman, A. and Zussman, N. (2006). Assassinations: Evaluating the effectiveness of an israeli counterterrorism policy using stock market data. *Journal of Economic Perspectives*, 20(2):193–206.