The London School of Economics and Political Science

Sustainable Development Challenges: Global Evidence and Case Studies in Vietnam, Indonesia, and Nigeria

Mukund (Mook) Bangalore

A thesis submitted to the Department of Geography and Environment of the London School of Economics and Political Science for the degree of Doctor of Philosophy, London, September 2023

Declaration

I certify that the thesis I have presented for examination for the MPhil/PhD degree of the London School of Economics and Political Science is solely my own work other than where I have clearly indicated that it is the work of others (in which case the extent of any work carried out jointly by me and any other person is clearly identified in it).

I warrant that this authorisation does not, to the best of my belief, infringe the rights of any third party.

I confirm that Chapter 1 was jointly co-authored with Dr Martin Heger and Dr Gregor Zens. For this chapter, I co-created the global dataset measuring socio-economic and environmental conditions, conducted statistical and geo-spatial analysis, and provided support in the write-up of the manuscript. I confirm that a version of this chapter is published in Environment and Development Economics, Volume 25, Issue 4, 2020.

I confirm that Chapter 2 was jointly co-authored with Dr Ulf Narloch. For this chapter, I conducted statistical and geospatial analyses and contributed to the write-up of the manuscript. I confirm that a version of this chapter is published in Environment and Development Economics, Volume 23, Issue 3, 2018. I can confirm parts of this chapter were the result of previous work I undertook while working at the World Bank.

I confirm that Chapter 3 was jointly co-authored with Dr Andrew Smith and Dr Ted Veldkamp. For this chapter, I led the overall research design, conducted empirical and geospatial analyses, and led the write-up the manuscript. I confirm that a version of this chapter is published in the Economics of Disasters and Climate Change, Volume 3, 2019. I can confirm parts of this chapter were the result of previous work I undertook while working at the World Bank.

I declare that my thesis consists of approximately 62,744 words.

Abstract

As a global community, we have many overarching goals to improve living standards and the state of our natural environment. To achieve these goals, examining the evolving relationship between human systems and environmental change at the global, national, sub-national, and local level is fundamental. This PhD thesis contributes to the academic and policy discourse and aims to better understand the relationship between environment and development across multiple scales. The empirical studies assembled provide evidence across The Global South and present case studies in Vietnam, Indonesia, and Nigeria on the dynamic relationship people have with environmental change and climate risk. Chapter 1 provides a broad overview of environment and development and investigates the link between poverty and land ecosystems at the global scale. Chapter 2 conducts an empirical case study of multiple sustainable development challenges in Vietnam. Chapter 3 also relates to Vietnam and focuses on one environmental risk which poses a challenge to communities today and in the future: flood risk and climate change. Chapter 4 also focuses on flood impacts and moves beyond exposure to examine vulnerability and the ability to respond of households in Nigeria. Chapter 5 examines how households respond to a policy-induced environmental change, exploring the impact of a cooking fuel conversion program in Indonesia. The evidence gathered through the five chapters of this PhD thesis further strengthen the call to integrate environmental considerations in economic development and poverty reduction policies. Despite the challenges and uncertainties that lie ahead in an era of climate change, opportunities exist to support the livelihoods of the poorest, protect the environment, and build more resilient economic systems.

Acknowledgements

The thesis may bear my name, but none of this would be possible without the support of my wonderful colleagues, family, and friends.

To my supervisors: Charlie – we've now known each other for more than a decade, both in the classroom and on the soccer pitch. Your constant support and guidance, through the ups and downs of the PhD journey both professionally and personally, has meant the world to me. Sefi – thanks for joining the team mid-way through, like a super-sub coming on at halftime. Your excellent comments and feedback has substantially improved the quality of my work. To my former supervisors – Antoine and Ben – your input has been extremely valuable to inform several chapters of this thesis, and I appreciate all the opportunities you've provided me over the years. I've learned so much, some (broken) French and soccer skills.

To my examiners: Diana and Jill: thank you very much for agreeing to serve on my panel and for taking the time and energy to help improve the quality of my work. Your contributions will help me become a better researcher. I am very grateful.

To my colleagues: To my talented co-authors Ulf, Martin, Andy, Ted, and Gregor: I am delighted to include the research papers we worked on together as part of this package. To the Grantham Research Institute: I am so grateful to have received a scholarship to be able to carry out this research. I've always loved being a part of the GRI community. Many others have provided research opportunities, data and contextual information to support the research, including Stephane, Harun, Keiko, and Janet.

To my family and friends: You've always been there with me through my trials and tribulations and supported me in so many ways. To Sanjana, my parents, my grandparents, my brother, my cousins, and more: I don't know where I'd be without all of you. But I do know I wouldn't be here now. Sometimes you may not understand my research, but you understand me. To my friends – your energy and spirit has brought the fire when the light was dim. There are so many to thank including: Lindsey, Janna, Anshuman, Matt, Andreas, Jared, Julius, Marta, Tobias U, Tobias K, Rajat, Xavi, Tom, Anirudh, David, Michael, Mike, Jun, Steve, Margarida, Matthais, Lutz, Joana, Misato, Gregor, Pierre, and so many others.

Table of Contents

List of Figures

List of Tables

Introduction

Adopted by the international community in 2015, the 2030 Agenda for Sustainable Development outlined 17 Sustainable Development Goals (SDGs) to improve living standards and protect the natural environment (Figure 1). To achieve these overarching goals, examining the evolving relationship between human systems and environmental change at the global, national, sub-national, and local level is fundamental. This PhD thesis contributes to the academic and policy discourse and aims to better understand the relationship between environment and development across multiple scales. The empirical studies assembled provide evidence across The Global South and present case studies in Vietnam, Indonesia, and Nigeria on the dynamic relationship people have with environmental change and climate risk.

Figure 1. The 17 Sustainable Development Goals (SDGs) outlined by the international community in 2015.

In The Global South, one of the major threats to achieving the SDGs is environmental change, which includes degradation to land, forests, and ecosystems, extreme weather events such as floods, droughts, and storms, as well as impacts from global climate change. Households in low and middle-income countries rely heavily on environmental capital for their livelihoods and lose a large share of their income and assets from degradation in environmental quality, the effects of weather shocks, and climate change impacts (Wunder, Noack and Angelsen, 2018; Hallegatte *et al.*, 2020; IPCC, 2023). In addition to environmental change, policy intervention at the local level can support the livelihoods of households, for example to promote the use of cleaner cooking methods. This PhD thesis explores household relationships with both environmental and policy changes, focusing on slow-run changes in land, forests, and multiple environmental categories across The Global South and specifically in Vietnam, extreme events from floods and climate change in Nigeria and Vietnam, as well as a policy to promote the adoption of clean cooking in Indonesia.

The thesis makes three main contributions to the existing literature. First, it provides descriptive and causal evidence at the micro-level on how environmental change impacts people's livelihoods across multiple countries, risk types, and time periods. Second, the thesis creates new databases to examine human-environment interactions and combines panel household survey data with geo-spatial information to investigate the dynamic relationships at a high-resolution of detail to uncover nuanced relationships and distributional impacts. Third, multiple scales are examined – the household, community, district, province, and region – which provides insights on the mechanisms that determine the relationship between environmental change and livelihoods. Now in 2023, halfway between adoption of the SDGs in 2015 and their expected completion in 2030, evidence from this PhD thesis can help inform the design and implementation of local policies to support livelihoods and protect the environment with the aim of achieving the overarching goals set by the international community.

The PhD thesis is comprised of 5 chapters (3 co-authored and 2 single-authored) which explore various sustainable development challenges across The Global South with case studies in Vietnam, Indonesia, and Nigeria. The structure is as follows.

Chapter 1 provides a broad overview of environment and development and investigates the link between poverty and land ecosystems at the global scale. The chapter is titled "Land and Poverty: The Role of Soil Fertility and Vegetation Quality in Poverty Reduction", is co-authored with Martin Heger and Gregor Zens, and was published in the journal *Environment and Development Economics*. In terms of the SDGs, the evidence from this paper relates primarily to SDG 15 (Life on Land), SDG10 (Reduced Inequalities), and SDG1 (No Poverty). This paper provides a global study across The Global South using quasi-experimental methods to uncover to what degree land improvements matter for poverty reduction.

Chapter 2 builds on the global analysis from Chapter 1 and conducts an empirical case study of multiple sustainable development challenges within one country: Vietnam. This chapter is titled "The multifaceted relationship between environmental risks and poverty: New Insights from Vietnam", is co-authored with Ulf Narloch, and was also published in the journal *Environment and Development Economics*. The evidence from this paper relates to SDG15 (Life on Land), SDG1 (No Poverty), SDG10 (Reduced Inequalities), and SDG3 (Good Health and Well-being). This paper combines geo-spatial data on eight environmental risks and household survey data from 2010-2014 to examine the complex relationship between environmental risk and poverty and how it differs across space, time, and scale within the same setting.

Chapter 3 also relates to Vietnam and focuses on one environmental risk which poses a challenge to communities today and in the future: flood risk and climate change. This chapter is titled "Exposure to floods, climate change, and poverty in Vietnam", is co-authored with Andrew Smith and Ted Veldkamp, and was published in the journal *Economics of Disasters and Climate Change.* Results from this paper relate primarily to SDG11 (Sustainable Cities and Communities), SDG1 (No Poverty), and SDG13 (Climate Action). The analyses examine the exposure to current and future climate change-related flooding across the country and specifically in Ho Chi Minh City, with a focus on poor communities.

Chapter 4 also focuses on flood impacts and moves beyond exposure to examine the vulnerability and ability to respond of households in response to a severe flood in Nigeria which occurred in 2012. This single-authored chapter is titled "Household exposure, vulnerability, and ability to respond to Nigeria's 2012 floods". In terms of the SDGs, this paper provides policy guidance for achieving SDG2 (Zero Hunger), SDG1 (No Poverty), and SDG13 (Climate Action). This paper examines the socio-economic and agricultural impacts of, and household responses to, an unprecedented flood which occurred in Nigeria in 2012.

In addition to impacts from environmental and climate change, Chapter 5 examines how households respond to government policies, exploring the impact of a cooking fuel conversion program in Indonesia. This single-authored chapter is titled "The impacts of fuel conversion on households: An assessment of Indonesia's Kerosene-to-LPG program in West Nusa Tenggara province". The results from this paper can help inform SDG7 (Affordable and Clean Energy), SDG3 (Good Health and Well-being), and SDG5 (Gender Equality). This paper examines the impacts of a policy to promote cleaner-burning LPG for cooking on socio-economic and health outcomes at the household level, with a particular focus on women and children.

Following the chapters, I conclude with a discussion of the external validity of the results, policy implications, and directions for future research.

Chapter 1: Land and Poverty: The Role of Soil Fertility and Vegetation Quality in Poverty Reduction (co-authored with Martin Heger and Gregor Zens)

A version of this article has been published as:

"Heger, M.P., Zens, G. and Bangalore, M., 2020. Land and poverty: the role of soil fertility and vegetation quality in poverty reduction. *Environment and Development Economics*, 25(4), pp.315-333."

Abstract: The debate on the land – poverty nexus is inconclusive, with past research unable to identify the causal dynamics. We use a unique global panel dataset that links survey and census derived poverty data with measures of land ecosystems at the sub-national level. Rainfall is used to overcome the endogeneity in the land-poverty relationship in an instrumental variable approach. This is the first global study using quasi-experimental methods to uncover to what degree land improvements matter for poverty reduction. We draw three main conclusions: First, land improvements are important for poverty reduction in rural areas and particularly so for Sub-Saharan Africa. Second, land improvements are pro-poor: poorer areas see larger poverty alleviation effects due to improvements in land. Finally, irrigation plays a major role in breaking the link between bad weather and negative impacts on the poor through reduced vegetation growth and soil fertility. However, it must be noted that within-province heterogeneity may also be driving the results and future research may want to exploit time variation within provinces or through a difference-in-difference design.

1.1. Introduction

The world had 1 billion fewer people living in poverty in 2013 compared to 1990 (measured in monetary terms (World Bank, 2016b). While poverty remains high, these aggregate numbers suggest that significant progress has been made in the past decades. Human capital formation, economic growth, trade, and institutional strengthening have been suggested as important drivers for this reduction in poverty headcounts (Ravallion, 2001; Bhagwati and Srinivasan, 2002; Harber, 2002). Economists often have a strong focus on these human development and macroeconomic drivers of poverty reduction (see e.g. Gennaioli *et al.*, 2013). Less emphasis has been placed on the role of the quality of renewable natural capital, such as healthy land ecosystems, which are the focus of this article. Notwithstanding, healthy land ecosystems – we will refer to them simply as "land" from here on forth, following convention (Nkonya, Mirzabaev and Von Braun, 2016) – are foundational for supporting livelihoods (e.g. (Angelsen *et al.*, 2014). According to Nkonya, Mirzabaev and Von Braun (2016), land improvement is closely approximated by two measures: net primary productivity (NPP) and soil fertility improvements. Hence, we focus on these two indicators in this article.

Early empirical studies have identified land degradation and declining soil fertility to be related to poverty at an aggregate level (Barrett and Swallow, 2006; Krishna *et al.*, 2006). More recently, Barrett and Bevis (2015) find that national GDP per capita is positively correlated with soil nutrient balances in 36 Sub-Saharan African countries for which data is available. Barbier and Hochard (2016) find that around a quarter of people living in low-income countries reside on severely degraded land and that a lower share of people on degraded land is associated with higher economic growth as well as lower poverty. Sanchez *et al.* (1997) stress the importance of soil quality for food security and development, especially in African countries. In addition, Koren (2018) finds a strong relationship between crop yields and conflict, which in turn is known to influence income and poverty (Goodhand, 2001).

Summarizing, the literature suggests a positive relationship between land quality and income. The theoretical channels behind this relationship are rather intuitive. One of the most important assets determining productivity for the rural poor is land (Barbier and Hochard, 2016, 2018). For instance, the water storage capability of soil is an important determinant of plant growth (Wong and Asseng, 2006). Louwagie *et al.*, (2011) find that shallow soils, stoniness or chemical issues such as salinity or acidity are negatively correlated with crop yields. In addition, the topographical conditions of the soil (elevation, steepness, etc.) affect soil erosion and accessibility by humans and machinery (e.g. Zuazo and Pleguezuelo, (2009). For an overview of the productivity function of soil, see Mueller *et al.* (2011). A main conclusion of the literature is that locations with good soils are likely to have high agricultural potential and thus have absolute advantage in generating agricultural income. However, it is not only increasing agricultural income that decreases poverty rates. Aside from crop cultivation and livestock income, forest management also reaps significant benefits that might alleviate poverty. Moreover, the so-called “hidden harvest’’ from the extraction of natural forests (i.e. forests that are not managed) and other non-forest wildlands as well as non-marketed extraction of natural resources can also play an important role in poverty reduction. For example, Angelsen *et al.* (2014) find that almost one-third of the total income of rural households is “environmental income,” of which more than three-fourths stem from natural forests.

Barrett and Bevis (2015) discuss three mechanisms through which poor land may have negative implications for poverty reduction. First, poor and degraded soils have negative effects on agricultural and environmental income. Such links can be self-reinforcing: poor soil constrains capital accumulation and low capital accumulation inhibits investments in improving soils (Eswaran *et al.*, 1997; Barrett and Bevis, 2015a). Second, poor and degraded soils are characterized by soil micronutrient deficiencies, which in turn can result in dietary mineral deficiencies affecting human health negatively (Barrett and Bevis, 2015a). The negative effect of deteriorating individual health on the ability to generate income is a long-standing fact in economics going back to Luft (1975). Third, low quality soils are connected to higher agricultural risks through various channels. For instance, weather shocks such as droughts occur more often in soils with limited water-holding capacity (Garrity *et al.*, 2010). In addition, pests and weeds, which decimate cropland in Sub-Saharan Africa, are more common in low-nutrient and degraded soils (Ayongwa *et al.*, 2011). There may be other mechanisms through which poor land affects poverty, such as through conflict. The link between conflict and land quality is much less understood however. Recent research has found significant effects between increasing spatial crop variability (within a country) and the probability of conflict (Ang and Gupta, 2018; Berman, Couttenier and Soubeyran, 2021). Other research has shown that higher yields are associated with more conflict (Koren, 2018).

Even though the theoretical implications are unambiguous, estimating the effect of healthy land (whether that is above-ground, as measured by improving vegetation quality, or below-ground as measured by improving soil fertility) on income and poverty is not a trivial issue for two reasons.

First, these variables are characterized by an endogenous relationship. Natural resources can influence poverty, but poverty can also influence natural resources (Barbier, 2010), with the relationship being moderated by economic, social and environmental factors (Barbier and Hochard, 2018). Due to potential simultaneity and intervening drivers, the causal effect of environmental quality on poverty reduction (and vice-versa) has been difficult to identify. As a result, most of the literature simply reports correlations (Duraiappah, 1998; Suich, Howe and Mace, 2015)[1]. Second, due to massive data collection efforts, prior studies are location-specific, and do not inform on how the relationship differs by biome or geographic region.

The main contribution of this paper is to provide quasi-experimental estimates of the impact of land improvements on poverty reduction[2]. Several additional contributions are presented: We use a global subnational dataset and monetary poverty rates that emerge from survey and census estimations rather than highly modelled measures (such as poverty measures derived from night lights or other satellite-derived information). This dataset is combined with measures of soil fertility and vegetation quality. This enables us to not only draw evidence from cross-sectional models, but also exploit variance over time by implementing a panel fixed effects model to minimize omitted variable bias. The presented findings are based on quasi-experimental identification of the effects of land on poverty. However, it must be noted that as we use country-fixed effects to explore the provincial relationship between land and poverty, we cannot rule out that within-province effects are driving the relationship of interest. With this limitation on internal validity in mind, our study has a relatively high degree of external validity due to the global scope of the analysis. In addition, instead of only analyzing the effect of cultivation income, we measure the effect of all land, not just land that is under agricultural use and management. This implies that other sources of income such as "environmental income" which is reaped in large parts from natural forests (Angelsen *et al.*, 2014) are also implicitly included in the analysis outlined below. With these methodological refinements, we obtain results that emphasize the importance of land for poverty reduction.

[1] A notable exception is Alix-Garcia, Sims and Yañez-Pagans (2015) who estimate the relationship for the case of Mexico.

[2] The quasi-experimental research design of this article is the main feature that distinguishes it from another recent contribution on the relationship between land degradation and poverty from Barbier & Hochard (2018).

1.2. Empirical Strategy

1.2.1. Data

We employ the Hidden Dimensions Dataset (HDD), a unique geospatial dataset linking environment and natural resource measures to poverty and other human development indicators at the subnational level, furnished by the World Bank. The geographical unit is the administrative unit 1 level, commonly referred to as "province" level. We use two different environmental measures: Net Primary Productivity (NPP) – our measure of above-ground land ecosystems, and Topsoil carbon content (soil fertility) – our measure of below-ground land ecosystems.

NPP is the rate at which an ecosystem accumulates biomass. It measures how much carbon dioxide plants take in during photosynthesis minus how much carbon dioxide is released during respiration. Hence, it is an indicator for how much of the absorbed carbon becomes part of leaves, roots, stalks or tree trunks. NPP data is captured via NASA's Terra and Aqua satellites. Generally, it has been found that NPP is a superior measure of biomass productivity and biodiversity (see for instance Phillips, Hansen and Flather, 2008) when compared to related indicators such as the Normalized Deviation Vegetation Index (NDVI). The average value per province is used for computations.[3]

Soil fertility is approximated by utilizing topsoil carbon content data. Topsoil carbon content is an important measure of plant productivity, measuring the percentage of carbon contained in the top 30cm of the soil. The carbon content of the soil is a result of e.g. decomposing plant and animal residues. It is a major determinant of plant growth and agricultural productivity (see for instance Lal, 2004). Hiederer and Köchy (2011) use the Harmonized World Soil Database to compute global soil organic carbon estimates on a subnational level. The Joint Research Centre of the European Commission provides this georeferenced dataset including both topsoil and subsoil carbon measurements via the European Soil Data Centre upon request.

The correlation of NPP and soil fertility may vary substantially. If the nutrient source for vegetation originates largely from the soils (i.e. soil based biomass productivity), then NPP is a very strong proxy of soil quality. If mineral fertilizers are used extensively, NPP is probably not a good indicator

[3] See the detailed documentation for the MODIS derived NPP measure at https://vip.arizona.edu/documents/MODIS/MODIS_VI_UsersGuide_June_2015_C6.pdf.

of soil quality (e.g. Nkonya, Mirzabaev and Von Braun (2016)). Hence, utilizing both measures in the empirical analysis is necessary to comprehensively analyze the concept of "land quality".

The measurement of poverty employed is the headcount ratio of people falling below $1.90 per day. Even though this is a narrow definition of poverty, $1.90 is the official international poverty line and allows us to draw from poverty maps that the World Bank produced for many countries over the last decades. This indicator captures what is commonly referred to as extreme poverty. From the World Bank poverty maps, a global map of sub-national poverty measures is created.[4] Gross domestic product per capita[5] is computed using GDP data from Gennaioli *et al.* (2013) and average annual population data from the Gridded Population of the World (GPW) dataset (CIESIN, 2016).

Mean average annual rainfall by province is used to instrument annual changes in vegetation quality and topsoil carbon. The data is sourced from the Climatic Research Unit in the National Center of Atmospheric Research (NCAR, 2017). The dataset contains geographically gridded multiple weather time series from 1901 onwards. The data is averaged by year and province for the computations. The source data is based on rain gauges. See Hulme (1992; 1994) and Hulme, Osborn and Johns (1998) for details on the spatial interpolation techniques employed to obtain a globally gridded dataset.

A number of variables are also included as controls. To capture the effect of different terrains, a topographic ruggedness index (Nunn and Puga, 2012) is used. This index captures small-scale terrain irregularities based on elevation differences. Land use categories (cropland, forest land, grass land, urban land and other) are included in the regression. Each land use indicator is measured as a share of the total geographic area. The original data is provided by the Land Cover project of the Climate Change Initiative led by the European Space Agency. In addition, a categorical variable corresponding to 14 different categories of soil types is included. These soil categories are a crucial determinant of soil quality. While a province may have several soil types, we assign the most

[4] The level of granularity of these poverty maps (most of them are at the province or admin 1 level) determines the granularity level of the analysis. It is the reason why the empirical analysis outlined below is based on the province level.

[5] Strictly speaking, we measure Gross Regional Product rather than Gross Domestic Product since our unit of analysis is not the country, but the province. However, as it is more conventional to refer to such economic activity as GDP, we stick to this nomenclature.

prevalent soil type to each province. This follows the soil classification system of the USDA system of soil taxonomy (USDA, 1999). Finally, we include road density and population as standard control variables.

In contrast to similar works, we do not rely on population data that is modelled using land use or night light data (e.g. Amaral *et al.*, 2005), which is commonly necessary for fine-grained spatial resolutions. Such datasets might raise severe endogeneity issues, as they could potentially be highly correlated with other satellite derived measures such as those measuring an environmental output. For instance, common measurement errors due to similarities of the satellites used to record the data could establish unwanted mechanical relationships in the data set. However, only the NPP measure is derived via remote sensing and earth observation, while the rainfall data is derived from rain gauges and statistical spatial interpolation. The employed poverty measures are based on censuses and surveys. Hence, we rule out the possibility of endogeneity that is an artifact of data construction or correlated measurement errors.

A description of all variables, including data sources, can be found in Table 1 and summary statistics are presented in Table 2. Overall, the data set contains 3303 observations for 1078 provinces in 62 countries. Coverage varies between 1996 and 2014, with country specific details provided in Table 1 of the Appendix A.

Table 1. Variable Overview

Variable	Units	Source
Poverty Headcount Rate ($ 1.90 PPP)	%	WB
Gross Regional Product	USD	Gennaioli et al., 2013
Net Primary Productivity	gC/m^2	NASA
Topsoil carbon content	tons per hectare	European Comission (JRC)
Soil classification	14 categories	NRDC
Share of area cropland	%	CCI
Share of area forest	%	CCI
Share of area grassland	%	CCI
Share of area urban	%	CCI
Ruggedness Index	Index (0 to 1,000,000)	Nunn & Puga, 2012
Road Density	mean of length in km per province	PBL GeoNetwork
Mean precipitation	millimeters per month	CRU
Population	Persons	GPW
Irrigation	% of total area irrigated	Global Irrigation Map v5

Table 2. Summary statistics of all variables used in the analysis

Variables	(1) N	(2) mean	(3) sd	(4) min	(5) p10	(6) p25	(7) p75	(8) p90	(9) max
Poverty headcount ratio	3,303	33.28	19.13	0.400	9.900	17.10	46.97	59.80	93.60
Share cropland	3,303	0.354	0.250	0	0.0545	0.157	0.538	0.734	0.967
Share forest	3,303	0.370	0.263	0	0.0127	0.127	0.572	0.734	0.977
Share grassland	3,303	0.0683	0.125	0	0	0.000500	0.0672	0.223	0.863
Share urban	3,303	0.0269	0.0998	0	0.000358	0.00143	0.0134	0.0376	0.955
Top soil carbon	3,297	64.87	51.76	7.998	25.42	31.77	74.58	142.4	318.1
Precipitation	2,802	4.557	0.837	-0.261	3.518	4.036	5.193	5.459	6.464
NPP	3,225	0.494	0.786	-5.081	-0.453	0.167	1.047	1.263	1.669
Road density	3,295	2.832	1.156	-2.316	1.532	2.145	3.584	4.151	6.548
Ruggedness	3,299	11.38	1.249	3.158	9.782	10.65	12.30	12.64	13.71
GDP per capita	2,287	-5.427	1.807	-11.52	-7.686	-6.572	-4.222	-3.391	4.860
Population	3,290	13.56	1.574	4.706	11.56	12.54	14.51	15.34	19.14
Irrigation	3,294	5.56	9.911	0	0.01	0.32	6.12	15.64	80.46

1.2.2. Research design

As stated previously, land shares a simultaneous relationship with income and poverty. There is a so-called poverty-degradation vicious cycle: poverty leads to degraded soils, while degraded soils lead to poverty (Eswaran *et al.*, 1997; Barbier, 2000; Lambin *et al.*, 2001). At the same time, it has also been found that reducing poverty rates can have positive or negative effects on degradation, depending on the initial levels of development (Cuaresma and Heger, 2019). Regardless of the direction of the bias, using standard OLS estimation would thus lead to biased coefficient estimates.

To overcome the methodological challenges arising from the endogenous relationship of land, income and poverty, a simultaneous equation model with instrumental variables is implemented. For vegetation quality, a panel regression is specified as time series are available for both NPP and rainfall data. However, the data for topsoil carbon is time-invariant. Therefore, a cross-sectional regression using the most recent observations per province is estimated. Similarly, the data set provides variation over time with respect to poverty, but regional GDPPC is measured at one point in time only, restricting us to cross-sectional specifications in these cases.

Panel specification of land quality and poverty

The panel regression to infer the effects of land quality on poverty is specified as below.

Equation 1. Panel regression of land quality on poverty

$$Y_{ijt} = \gamma_j + \delta_t + \alpha D_{ijt} + \beta X'_{ijt} + \varepsilon_{ijt}$$

Here, Y_{ijt} is the poverty headcount rate for province i in country j in year t, γ_j represents the country-fixed effects, δ_t represents time-fixed effects, αD_{ijt} is the main NPP explanatory variable for each province in each country at each time period, $\beta X'_{ijt}$ includes covariates (five land use categories, population, ruggedness index, and road density), while ε_{ijt} represents the error term. While the unit of analysis is the province level, country-fixed effects are used. When interpreting the results, it is important to keep in mind that within-country province heterogeneity may be driving the relationships of interest. For example, due to internal selection and sorting across provinces, poorer households may reside in provinces with lower land quality. Even though this issue is somewhat alleviated by including variables that partially account for some province heterogeneity, this constitutes a notable shortcoming of our analysis.

Cross-sectional specification of land quality, poverty and income

Equation 1 identifies the effect of the environment exploiting variation over time. However, GDP per capita and topsoil carbon content data are time-invariant. Thus, the effects of soil fertility can only be assessed using spatial variation. The same holds true for the effects of vegetation quality on GDP.

Equation 2. Cross-sectional regression of land quality on income/poverty

$$Y_{ij} = \gamma_j + \delta_t + \alpha D_{ij} + \beta X'_{ij} + \varepsilon_{ij}$$

Here, Y_{ij} is either the poverty headcount rate or GDP per capita for province i in country j, γ_j represents the country-fixed effects, δ_t represents time-fixed effects[6], αD_{ij} is the main topsoil carbon explanatory variable for each province in each country, $\beta X'_{ij}$ includes covariates (five land use categories, population, ruggedness index, and road density), while ε_{ij} represents the error term. Robust standard errors are used in all specifications. NPP, GDP per capita, top soil carbon, population, ruggedness, road density and precipitation enter the model after a log-transformation

[6] Time fixed effects capture the fact that the observations in the model stem from different years in this case.

as guided by the existing literature (Nunn and Puga, 2012; Gennaioli *et al.*, 2013; Cuaresma and Heger, 2019). As the data is highly skewed, the log-transformation reduces the likelihood outliers are driving the main results and aids the interpretability and usefulness of the estimates when inferred through percentage change variations (Cuaresma and Heger, 2019; World Bank, 2023a).

Instrumental variables approach

Instrumental variable estimation is employed to overcome the endogeneity between vegetation quality and soil fertility and poverty. Rainfall is used as a source of exogenous variation for NPP and topsoil carbon. Nevertheless, there has been a debate on the validity of rainfall as an external instrument. After careful review of the pertinent literature, we conclude that rainfall is a viable instrument for our research design as rainfall is a strong determinant of above- and below-ground biomass, meets the exclusion criterion, and is as-if randomly assigned. More details are discussed below.

Rainfall is one of the most crucial determinants of vegetation quality and biomass productivity (for evidence, see pertinent agronomic literature such as Vlam *et al.* (2014) and Schippers *et al.*, (2015)). Precipitation influences soil moisture and above-ground biomass by affecting seed germination, seedling growth, and plant phenology (Kang *et al.*, 2013; Liu *et al.*, 2014; Yan *et al.*, 2014). Furthermore, precipitation is also the main input factor for soil fertility: the greater the biomass production resulting from more rainfall, the more residues are produced, which in turn leads to more potential food for soil biotas. Testament to the major importance of rainfall for soil fertility (and in particular for soil organic carbon) is the fact that precipitation is the main input factor in Revised Universal Soil Loss Equation models (see e.g. (Angulo-Martínez and Beguería, 2009; Hernando and Romana, 2015)) and in the GIS-based Universal Soil Loss model (Angima *et al.*, 2003; Lufafa *et al.*, 2003; Fu *et al.*, 2005).

We argue that if there is a fitting case for using rainfall as an IV, using it for isolating the exogenous variation in vegetation and soil quality is one of the most promising candidates. Rainfall is extremely closely linked to the treatment variables (vegetation growth and soil fertility) in our study.[7] In fact, rainfall is perhaps the most important determinant of plant growth, particularly so in areas with

[7] This becomes obvious from the first stage regressions in the results section below.

little irrigation.[8] The economies of low-income and middle-income areas are particularly dependent on the primary sector such as agriculture and forestry. Increased quantities of rainfall increase crop yields and the environmental income from surrounding ecosystems, a mechanism which ought to be especially strong in Sub-Saharan Africa, where only 4% of area cultivated is equipped for irrigation as compared to for instance 28% in North Africa (You *et al.*, 2011).

An influential article by Sarsons (2015) casts doubt as to the validity of rainfall as an instrument for conflict. She showed that in irrigated areas, rainfall shocks are a weaker predictor for income changes, but nevertheless remains a significant predictor of conflict, indicating that there are channels other than income through which rainfall affects conflict.[9] Note that this criticism does not directly apply to our research design, as we use rainfall as an instrument for soil fertility and vegetation quality (and then investigate its effects on income). However, her larger point remains also a valid criticism to our identification strategy, as she suggests that income may be affected by rainfall through channels outside of agricultural, forestry, and other environmental reasons.

One major concern with using rainfall as an IV for income (poverty) is that extreme rainfall events (such as flooding) can lead to the destruction of property and affect poverty outside of the channels of soil fertility and vegetation changes, therefore violating the exclusion restriction. For example, floods may affect transportation and the ability to organize. A similar concern applies to droughts, which might kill livestock due to heat stress. We overcome the flooding and drought identification threat by excluding outlier rainfall events in separate specifications below.[10] Furthermore, by including road density and ruggedness we control for the transportation identification threat.[11] A suggestive empirical indication that this exclusion restriction holds, is that the OLS specifications below indicate that rainfall is not a statistically significant predictor of poverty rates or GDP per capita when controlling for environmental quality.

---

[8] We specifically look at the importance of irrigation for the environment-poverty elasticity in the results section.
[9] This furthermore suggests that the exclusion restriction in several seminal papers, including e.g. Paxson (1992), Miguel, Satyanath and Sergenti (2004), Miguel (2005) and Yang and Choi (2007) may be violated.
[10] For this, we exclude the top and bottom 10% of rainfall events from the sample.
[11] There may be other channels through which rainfall may affect welfare which we have not explored, as they have not (yet) been discussed in the literature. However, that may be said of any IV.

Other identification compromising channels that Sarsons (2015) describes are migration, where farmers move from rain-fed to dam-fed provinces, creating a conflict over scarce land. She also describes spillover effects as another channel that may violate the exclusion restriction. Her point is that violence may propagate from a violent rain-fed to an initially non-violent dam-fed province, explaining why rainfall also affects violence in dam-fed provinces. This criticism may also extend to using income as an outcome variable, as for example, conflict also affects income (Blattman and Miguel, 2010).

To overcome this issue, we split our sample based on irrigation to separately analyze relatively well irrigated and relatively badly irrigated areas. Similarly, Sarsons (2015) discussed dam-fed provinces and rain-fed provinces separately. Note that Sarsons (2015) shows that conflict is affected by rainfall regardless of irrigation as evidence for a violated exclusion restriction. On the contrary, we find that irrigation actually explains a significant proportion of the environment-poverty elasticity. This points in the direction of an upholding exclusion restriction.

Finally, it is worth mentioning that rainfall is randomly assigned as weather is an exogenous event in each province. Even if climate change, which is clearly affected by development, alters rainfall patterns, it does so on a global scale, and it is hardly attributable to a given province's actions alone, therefore not violating the as-if random assignment assumption.

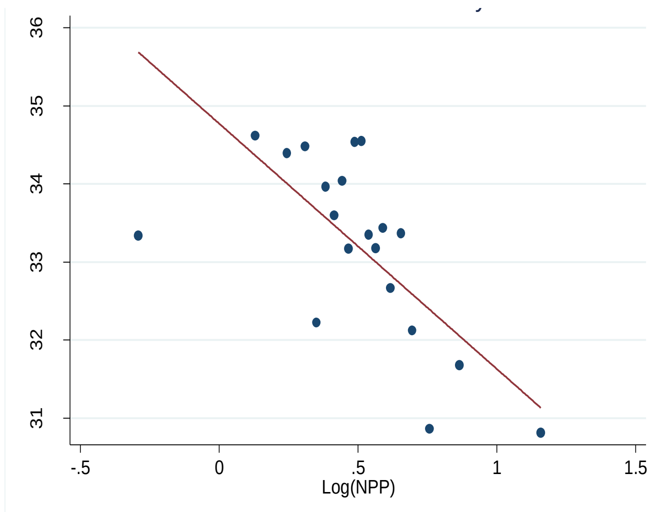
1.3. Results

We detail three main findings: First, vegetation quality and soil fertility have significant and sizeable poverty alleviating effects, particularly in rural areas and especially so in Sub-Saharan Africa. Second, improving soil & vegetation quality is pro-poor: poverty rates in areas with high poverty headcounts are significantly more strongly affected by improvements in soil and vegetation than areas with relatively low poverty incidence. Finally, the dependence of rural areas on rainfall-induced changes in vegetation quality and soil fertility is reduced by irrigation. We thus conclude that irrigation systems have significant impacts for making the poor resilient to the vagaries of weather and climate. However, when interpreting the all the results in this section, one threat to internal validity has to do with within-province heterogeneity. As country-level fixed effects are used (instead of province-level fixed effects), time invariant spatial correlation across provinces between land quality and poverty may be driving the relationship of interest and is not fixed by the instrumental variable estimation. With this important caveat in mind, the main results are presented below.

1.3.1. Panel Fixed Effects evidence

Figure 2 shows a strong correlation between vegetation quality and poverty. The graph shows the conditional relationship of NPP and poverty after controlling for other possible predictors of poverty. It seems obvious that increasing vegetation quality is associated with accelerated poverty reduction.

Figure 2. The relationship between changes in vegetation & poverty reduction



Note: Each dot represents an equally sized bin of observations (grouped over the x-axis). Within these bins, the average of the x- and y-variable is computed and visualized in a scatterplot. The plot gives the conditional effect of the natural logarithm of NPP on the residualized poverty headcount ratio after controlling for several covariates. They are created by running an OLS regression equivalent to Table 5, column (1).

However, from this descriptive analysis it does not automatically follow that vegetation quality causally influences poverty reduction for the average province in our sample. Table 3 and Table 4 outline results of the instrumental variables strategy. In general, the first-stage regressions show a very strong relationship of vegetation quality and poverty: precipitation explains more than 80% of the variation of NPP as shown in Table 3.

Table 4 provides results on the second-stage regressions and shows that despite the OLS specification (1) being significant, the global IV specification (4) is not. However, vegetation quality seems to be much more important for more rural areas, as one would expect, as seen in specification (5) and (6).[12] The panel results show that an increase of vegetation quality (NPP) by ten percent in rural areas reduces poverty rates by around 0.7 percentage points. In Sub-Saharan Africa, the effects were even larger, such that a ten percent increase in NPP resulted in a 1.2 percentage point decrease in poverty rates. The reasons for such significant and sizeable effects in Sub-Saharan Africa and rural areas has likely to do with livelihoods there being comparably more dependent on vegetation quality and soil fertility. For instance, Barrios, Bertinelli and Strobl (2010) have shown that unlike in other continents, economic growth is strongly dependent on rainfall in Africa. Moreover, Alene *et al.* (2018) find that soil fertility management had the largest effect on poverty reduction and economic growth in Africa. However, the estimates for rural Sub-Saharan Africa can be interpreted cautiously as the comparably low sample size results in statistical limitations with respect to inferring the exact effect size. Introducing controls makes the analysis even less statistically powered, which is the reason why the rural SSA specifications are in fact excluded from the specifications below.

Table 3. First stage – the effect of rainfall on NPP

Variables	(1) IV	(2) IV - Rural	(3) IV - Rural no outliers	(4) IV - Rural SSA
Log Precipitation	0.62***	0.63***	0.64***	1.74***
	(0.04)	(0.07)	(0.08)	(0.19)
Constant	-3.57***	-3.64***	-3.66***	-8.37***
	(0.18)	(0.56)	(0.57)	(0.92)
Observations	2,738	1,362	1,306	104
R-squared	0.81	0.81	0.80	0.64
Controls	YES	YES	YES	YES
Country FE	YES	YES	YES	NO
Year FE	YES	YES	YES	NO

Notes: All control variables are included. SSA stands for sub-Saharan Africa, Rural indicates the analysis is run in rural provinces only, and “No Outliers” excludes outlier rainfall events. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 4. Second Stage - The effect of NPP on Poverty

VARIABLES	(1) OLS	(2) OLS - Rural	(3) OLS - SSA	(4) IV	(5) IV - Rural	(6) IV - Rural No Outliers	(7) IV - Rural SSA
Log NPP	-1.54	-0.95	0.05	0.13	-7.20***	-7.20***	-12.14***
	(1.20)	(1.28)	(1.26)	(1.54)	(2.40)	(2.48)	(2.80)

[12] We define provinces with a crop share above 30% as ‘‘rural“ to simultaneously capture high levels of agricultural dependence and a low degree of urbanization.

Log NPP * Rural		-1.54					
		(1.16)					
Log NPP * SSA			-9.24***				
			(2.04)				
Log Precip	1.03	0.86	0.67				
	(1.25)	(1.24)	(1.26)				
Constant	39.21 ***	40.62***	43.17***	33.37***	31.15***	31.35***	41.91
	(6.18)	(6.20)	(6.32)	(7.21)	(8.17)	(8.18)	(26.01)
Observations	2,738	2,738	2,738	2,738	1,362	1,306	104
R-squared	0.68	0.68	0.69	0.68	0.73	0.73	0.52
Country FE	YES	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES	YES

Notes: NPP stands for net primary productivity and is log-transformed. Precip is short for precipitation and is also log-transformed. SSA stands for sub-Saharan Africa, Rural indicates the analysis is run in rural provinces only, and "No Outliers" excludes outlier rainfall events. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 5 shows a specification including a set of additional control variables. All estimated coefficients are in line with our priors. Note that all effects other than those of vegetation quality are not 'exogenized', thus they may not be interpreted as causal. That said, important lessons can be drawn from correlations as well. For example, road density, a proxy for infrastructural development, is an important and strong predictor of poverty reduction. This a long established and well-known finding in development economics (Jacoby, 2000; Gibson and Rozelle, 2003; Jacoby and Minten, 2009; Khandker, Bakht and Koolwal, 2009). For an overview of possible theoretical channels see Brenneman and Kerf (2002). Similar effects can be seen with respect to urbanization (people moving to cities), another variable that is consistently correlated with poverty reduction in the development economics literature (Christiaensen, De Weerdt and Todo, 2013; Arouri, Youssef and Nguyen, 2017). Ruggedness is statistically positively related to poverty rates, suggesting that rougher terrains possibly make it harder to escape poverty. The direction of control variable coefficients is generally in line with the literature, which gives confidence in the quality of the data and empirical approach.

Table 5. Second Stage - The effect of NPP on Poverty including Controls

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	OLS	OLS - Rural	OLS - SSA	IV	IV - Rural	IV - Rural No Outliers
Log Precipitation	1.52	1.51	1.33			
	(1.32)	(1.32)	(1.33)			
Log NPP	-4.18***	-3.47***	-2.66*	-0.96	-6.70**	-5.79**
	(1.21)	(1.33)	(1.36)	(2.47)	(2.76)	(2.70)
Log NPP * Rural		-1.30				
		(0.99)				
Log NPP * SSA			-6.19***			

			(2.18)			
Share Cropland	9.78***	9.85***	7.94**	8.62**	6.75	6.32
	(3.48)	(3.49)	(3.54)	(3.84)	(5.40)	(5.72)
Share Urban	-0.54	-1.19	-1.38	0.21	-25.53***	-27.11***
	(7.01)	(7.10)	(7.19)	(7.22)	(9.62)	(9.96)
Share Grassland	-1.58	-2.38	-2.98	-1.70	13.78	13.86
	(5.33)	(5.41)	(5.40)	(5.33)	(10.38)	(10.67)
Share Forest	3.67	2.23	1.67	0.56	0.07	-0.84
	(4.10)	(4.26)	(4.18)	(5.43)	(7.23)	(7.11)
Log Population	-1.85***	-1.90***	-1.81***	-1.82***	-1.10**	-1.14**
	(0.57)	(0.57)	(0.57)	(0.55)	(0.55)	(0.54)
Log Ruggedness	1.87***	1.92***	1.67***	1.55***	3.02***	2.94***
	(0.42)	(0.42)	(0.42)	(0.48)	(0.55)	(0.53)
Log Road Density	-2.90***	-2.90***	-2.87***	-3.09***	-5.02***	-5.08***
	(0.80)	(0.81)	(0.82)	(0.83)	(1.23)	(1.24)
Constant	47.60***	49.01***	53.66***	42.06***	14.43	17.09
	(10.13)	(10.16)	(10.48)	(13.30)	(12.65)	(12.23)
Observations	2,736	2,736	2,736	2,736	1,360	1,304
R-squared	0.74	0.74	0.75	0.74	0.82	0.82
Country FE	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES
Country trend	YES	YES	YES	YES	YES	YES

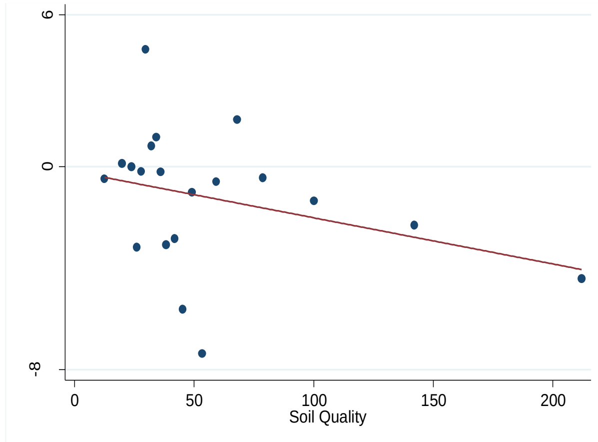
Notes: NPP stands for net primary productivity and is log-transformed (as are precipitation, population, ruggedness, and road density). SSA stands for sub-Saharan Africa, Rural indicates the analysis is run in rural provinces only, and "No Outliers" excludes outlier rainfall events. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

The main caveat of the presented analysis is that we are not able to control for province level fixed effects due to data availability, as elaborated in footnote 6. Not being able to control for province level fixed effects implies that we cannot rule out that the presented results are partially due to within country province heterogeneity. While including several control variables allows us to rule out the impact of several important within-country factors such as land cover and land use, it does not allow us to rule out any possible effect that would come from time-invariant within-country heterogeneity.

1.3.2. Cross-sectional findings

Figure 3 depicts a negative relationship between poverty and soil quality. Poverty headcount rates are particularly high in rural regions with low soil quality.

Figure 3. Poverty and Soil Quality in Rural Areas



Note: Each dot represents an equally sized bin of observations (grouped over the x-axis). Within these bins, the average of the x- and y-variable is computed and visualized in a scatterplot. The residuals from a regression of poverty headcount ratio on country fixed effects are on the y-axis. Top soil carbon content is on the x-axis.

Similar to the first-stage regressions of precipitation on NPP, the first-stage regressions show a significant relationship between precipitation and top-soil carbon (Table 6). In terms of the second-stage, the negative relationship between soil fertility and poverty is significant in the specifications that isolate the exogenous effects of top soil on poverty (Table 7). An increase in top soil carbon content of ten percent reduces the poverty headcount ratio by around two to three percentage points (columns 3 and 4) in the rural sample. The effects are especially large in rural Sub-Saharan Africa, where a ten percent increase in soil fertility results in a roughly four percentage points reduction in poverty rates in the baseline specification (column 5). As mentioned before, these specifications are restricted to cross-section information due to the unavailability of time-variant soil fertility and subnational GDP measures.

Table 6. First stage – the effect of precipitation on top soil carbon

	(1)	(2)	(3)	(4)
Variables	IV	IV - Rural	IV - Rural no outliers	IV - Rural SSA
Log Precipitation	0.27***	0.20***	0.25***	0.34***
	(0.03)	(0.04)	(0.05)	(0.08)
Constant	2.03***	2.18***	1.98***	1.90***
	(0.13)	(0.19)	(0.22)	(0.42)
Observations	933	476	452	64
R-squared	0.74	0.74	0.71	0.65
Controls	YES	YES	YES	YES
Country FE	YES	YES	YES	NO

Year FE	YES	YES	YES	YES

Notes: All control variables are included. SSA stands for sub-Saharan Africa, Rural indicates the analysis is run in rural provinces only, and "No Outliers" excludes outlier rainfall events. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 7. Second Stage - The effect of Top Soil Carbon on Poverty

	(1)	(2)	(3)	(4)	(5)
VARIABLES	OLS	IV	IV - Rural	IV - Rural No Outliers	IV - Rural SSA
Log Top Soil Carbon	-4.17***	0.73	-28.62***	-21.72***	-40.74***
	(1.01)	(3.23)	(9.32)	(8.11)	(11.25)
Log Precipitation	1.32				
	(0.94)				
Constant	47.34***	35.10***	134.72***	115.02***	212.56***
	(5.88)	(10.50)	(26.76)	(23.33)	(40.85)
Observations	933	933	476	452	64
R-squared	0.75	0.74	0.71	0.76	0.51
Country FE	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES

Notes: Results with and without controls are very similar, and for brevity the results above are without controls. Full results with controls can be found in Table 2 of Appendix A. Top soil carbon and precipitation are log-transformed. SSA stands for sub-Saharan Africa, Rural indicates the analysis is run in rural provinces only, and "No Outliers" excludes outlier rainfall events. Robust standard errors in parentheses.. *** p<0.01, ** p<0.05, * p<0.1

The significant results we estimate for the effect of soil carbon on poverty are robust with respect to the specific welfare measure chosen. We repeat above analysis using GDP per capita as the outcome variable instead of poverty headcount rates and find that similar patterns hold. An increase of topsoil carbon by ten percent results in an increased GDP per capita of 0.2 percent in rural areas (see columns 3 and 4 of Table 8). The first stage regressions estimating the effect of precipitation on top soil carbon are provided in Table in the Appendix A.

Table 8. Second Stage – The effect of Top Soil Carbon on GDP per capita

	(1)	(2)	(3)	(4)	(5)
VARIABLES	OLS	IV	IV- rural	IV - Rural No Outliers	IV - Rural SSA
Log Top Soil Carbon	0.23**	0.17	1.78**	1.68*	1.26***
	(0.11)	(0.30)	(0.88)	(0.90)	(0.37)
Log Precipitation	-0.02				
	(0.12)				
Constant	-3.76***	-7.03***	-13.25***	-12.91***	-12.06***
	(0.63)	(1.13)	(3.36)	(3.44)	(1.27)
Observations	636	636	339	330	32
R-squared	0.79	0.79	0.83	0.83	0.24
Country FE	YES	YES	YES	YES	NO
Year FE	YES	YES	YES	YES	NO

Notes: Top soil carbon and precipitation are log-transformed. SSA stands for sub-Saharan Africa, Rural indicates the analysis is run in rural provinces only, and "No Outliers" excludes outlier rainfall events. Robust standard errors in parentheses. Results with and without controls are very similar, and for brevity the results above are without controls. Full results with controls can be found in Table 3 of Appendix A. *** p<0.01, ** p<0.05, * p<0.1.

The significant effects of soil fertility on poverty and GDP per capita are moreover robust to the inclusion of the set of control variables already discussed in the panel specification for NPP. The full specifications including land use categories, road density and ruggedness are found in Table 2 and Table 3 in the Appendix A. In addition, we included a measure of soil type as a control variable, which is particularly important as there are significant variations of soil types within countries. Soil types range from soils with relatively rich soil organic carbon (such as Histosols) to soils with practically no soil carbon (such as Entisols). Clearly it is important to control for the specific type of soil as this is one of the main immutable factors when it comes to soil organic carbon formation (for a taxonomy of soils and an exposition of their properties see (USDA, 1999).

1.3.3. Additional findings

As discussed earlier, irrigation systems in Sub-Saharan Africa are much less developed than elsewhere, making the region more vulnerable to the vagaries of rainfall (You *et al.*, 2011). We therefore further investigate the role of irrigation and its effects on the environment-poverty elasticity, directly. Following the criticism brought forward in Sarsons (2015), we run split sample regressions based on irrigation prevalence in the provinces under analysis. The results are shown in Table 9. The estimates suggest that the effects of vegetation quality on poverty are indeed driven by the less irrigated areas in the sample. This suggests that irrigation systems are effective in increasing rural farmer's resilience to weather shocks.[13]

In an additional exercise, we analyze the degree of poverty alleviation along the income distribution using a split sample regression with areas below and above the observed median poverty rate. Table 10 and Table 11 show that for both the NPP and the soil organic carbon specifications, the

[13] That said, it is important to note that the sample split based on irrigation prevalence, may have split the sample also along the lines of several omitted variables. For example, the reason for better irrigation in one province, compared to the other may have something to do with quality of governance (see e.g. (Playán, Sagardoy and Castillo, 2018), which in turn may have mediated the strength of the environment-poverty elasticity, rather than irrigation per se. Future research into this area is necessary.

poverty rates in poorer places dropped much more as a reaction to improved vegetation quality and soil fertility. This is indicated by estimated coefficients that are in the order of two to nine times larger than the coefficient for less poor areas.

Table 9. The Effect of NPP on Poverty in Areas with High and Low Levels of Irrigation

	(1)	(2)	(3)
VARIABLES	IV - Rural	IV - Rural Above Med. Irrig.	IV - Below Med. Irrig.
Log NPP	-7.20***	-4.04	-21.69***
	(2.40)	(3.36)	(8.25)
Constant	31.15***	38.36***	38.43***
	(8.17)	(8.32)	(12.51)
Observations	1,362	855	457
R-squared	0.73	0.77	0.71
Country FE	YES	YES	YES
Year FE	YES	YES	YES

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Note: The sample is split according to the average proportion of area equipped for irrigation to area according to version 5 of the Global Map of Irrigation Areas published by the Food and Agriculture Organization of the United Nations.

Table 10. Second Stage – The effect of NPP on Poverty in High / Low Poverty Areas

	(1)	(2)	(3)
VARIABLES	IV - Rural	IV - Rural Below Median Pov.	IV - Rural Above Median Pov.
Log NPP	-7.20***	-2.50	-6.92***
	(2.40)	(2.43)	(2.56)
Constant	31.15***	3.46***	24.47***
	(8.17)	(1.33)	(8.38)
Observations	1,362	728	634
R-squared	0.73	0.44	0.56
Country FE	YES	YES	YES
Year FE	YES	YES	YES

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 11. Second Stage – The effect of Top Soil Carbon on Poverty in High / Low Poverty Areas

	(1)	(2)	(3)
VARIABLES	IV - Rural	IV - Rural Below Med. Pov.	IV - Rural Above Med. Pov.
Log Top Soil Carbon	-28.62***	-3.64	-32.99**
	(9.32)	(7.69)	(13.44)
Constant	134.72***	27.32	143.86***
	(26.76)	(30.59)	(39.93)
Observations	476	283	193
R-squared	0.71	0.58	0.50
Country FE	YES	YES	YES
Year FE	YES	YES	YES

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

To account for the possibility that the results could be driven by one explicit measurement of vegetation quality, all specifications are estimated with a different satellite-based indicator, the Normalized Density Vegetation Index. In addition, some spatial econometrics exercises are carried out. We specify spatial autoregressive models to account for the possibility of spatial dependence of poverty and income. In addition, we run Moran's I tests on the residuals of the second stage regressions of selected IV specifications. The conclusions drawn remain unchanged.[14]

1.4. Conclusion and discussion

In this article, we analyze the relationship of vegetation quality and soil fertility with income and poverty on a global scale. To overcome potential endogeneity issues, the exogenous variation of two environmental variables is isolated using rainfall data in an instrumental variable approach. However, it must be noted that province-level fixed effects are not included in the analysis, which are a limitation of this study and could potentially be driving the relationship of interest. Future research may want to examine a difference-in-difference type research design, with treated provinces as those potentially exposed to positive rainfall shocks. Efforts to improve the frequency of data collection at the provincial level on poverty headcounts and consumption may be a priority to allow researchers to examine a longer time series and explore the relationship between weather, land, and poverty. For example, newer methods to calculate poverty levels through phone surveys examine a longer time series and explore the relationship between weather, land, and poverty. For example, newer methods to calculate poverty levels through phone surveys can potentially provide higher-frequency data (Kakietek *et al.*, 2022). An additional limitation of the analysis is that we do not control for the total area of the province, as administrative areas can change over time. For example, provinces can become smaller in size as regions become richer and more populated. While increased wealth may reduce the province size and also be associated with changes in NPP and topsoil carbon, it may not lead to reverse causality as we are only focusing on precipitation induced changes of NPP and topsoil carbon on socio-economic outcomes.

[14] The NDVI results as well as the results of the spatial models are available upon request.

With these caveats in mind, we find evidence that building roads, and urbanization is associated with reductions in poverty rates, as previous literature suggested. What has not been shown conclusively so far is whether in-situ improvements of environmental quality significantly reduce poverty. Several authors have concluded that it does not. Okwi *et al.* (2007) conclude that if all of Kenya's soil was raised to its highest quality, only a one percentage point reduction in poverty rates would ensue. Wantchekon and Stanig (2015) go even farther and conclude that in Sub-Sharan Africa good soil may be a hindrance for poverty reduction.

We find that vegetation quality and soil fertility are important drivers for poverty alleviation in rural areas and Sub-Saharan Africa. Soil fertility and vegetation quality not only have significant and sizeable effects on poverty rates but also on GDP per capita. These significant environment-poverty elasticities are especially relevant for low income households that draw a larger share of their income from natural resources and the environment (Wunder, 2015). Moreover, we found that the effects of vegetation quality and soil fertility on poverty are stronger for poorer places, suggesting that environmental improvements are pro-poor. Finally, the availability of irrigation systems plays a major role when explaining the environment – poverty nexus. The results in this article are strongly driven by less irrigated areas, suggesting that irrigation systems have large impacts for making poor areas less dependent on weather fluctuations.

Chapter 2: The multi-faceted relationship between environmental risks and poverty: new insights from Vietnam (co-authored with Ulf Narloch)

A version of this article has been published as:

“Narloch, U. and Bangalore, M., 2018. The multifaceted relationship between environmental risks and poverty: new insights from Vietnam. *Environment and Development Economics*, 23(3), pp.298-327.”

Abstract: Despite complex interlinkages, insights into the multifaceted relationship between environmental risks and poverty can be gained through an analysis of different risks across space, time and scale within a single context using consistent methods. Combining geo-spatial data on eight environmental risks and household survey data from 2010-2014 for the case study of Vietnam, this paper shows: (i) at district-level the incidence of poverty is higher in high risk areas, (ii) at household-level poorer households face higher environmental risks, (iii) for some risks the relationship with household-level consumption varies between rural and urban areas, and (iv) environmental risks explain consumption differences between households, but less so changes over time. While altogether these analyses cannot establish a causal relationship between environmental risks and poverty, they do indicate that Vietnam’s poor are disproportionally exposed. Given growing pressures due to climate change, addressing such risks should be a focus of poverty reduction efforts.

2.1. Introduction

When devising poverty reduction strategies in the face of climate change, the role of environmental risks in the livelihoods of poor people needs to be understood. First, people struggling with multiple environmental risks generally have a lower ability to cope with other shocks and are thus more vulnerable to climate change impacts. Second, climate change will directly exacerbate weather-related environmental risks, such as rainfall and temperature variability and extremes and thereby increase this baseline vulnerability. Accordingly, poor people already exposed to high environmental risks could be most affected by climate change.

The interlinkages between poverty and environmental risks have been reviewed in earlier papers covering the extensive literature on this topic (Reardon and Vosti, 1995; Duraiappah, 1998; Scherr, 2000; Gray and Moseley, 2005; Barbier, 2010, 2012; Barrett, Travis and Dasgupta, 2011). Environmental risks are unevenly distributed across space – depending on geographic and climatic conditions, as well as socioeconomic factors that condition them. Poor people often live in remote and fragile areas, with high levels of environmental risks, where they overly depend on the use of ecosystems and natural resources, which can increase ecosystem fragility and environmental risks. This downward spiral could result in spatial poverty traps, where pockets of fragility, risks, marginalization and poverty persist (Carter *et al.*, 2007; Barbier, 2010; Barrett, Travis and Dasgupta, 2011). These poverty traps could be further exacerbated by climate change.

To guide policies to break out of these traps, there is a need for more in-depth, spatial analyses trying to disentangle the relationship between environmental risks and poverty (Barbier, 2012). While there is ample empirical work showing a positive relationship between environmental risks and poverty at the global and regional level (e.g., Chomitz *et al.*, 2007; Barbier, 2010, 2015; Sloan and Sayer, 2015) national and sub-national level (e.g., Dasgupta *et al.*, 2005; Bandyopadhyay, Shyamsundar and Baccini, 2011; Winsemius *et al.*, 2018) or in specific locations (Khan and Khan, 2009; Watmough *et al.*, 2016), all this work together points at a complex picture, where the relationship depends on the specific types of risks and locations of interest, as well as the channels through which they interact. New, high-resolution spatial data, for example from remote sensing, allows to analyze the distribution of environmental risks and poverty across space to better understand the multifaceted nature of this relationship.

Vietnam provides an interesting case study for such an analysis. Despite a reduction of extreme poverty from around 60 percent in 1990 to 13.5 percent as of 2014 (World Bank, 2016b),[15] pockets of poverty still exist today: 91 percent of people in extreme poverty live in rural areas and belong to marginal groups, such as ethnic minorities, who have a poverty rate of about 58 percent (Kozel, 2014). At the same time Vietnam is facing various environmental challenges. Land is increasingly degraded by human-induced factors (Vu *et al.*, 2014; Vu, Le and Vlek, 2014). Although total tree cover has increased, certain areas suffer from high rates of forest loss and degradation (Pham *et al.*, 2012). Air pollution is an increasing concern (Luong *et al.*, 2017). And weather variability and extreme events, such as floods and droughts, severely affect livelihoods (Thomas *et al.*, 2010; Bui *et al.*, 2014; Arouri, Nguyen and Youssef, 2015; Narloch, 2016). An earlier spatial analysis at the provincial level in Vietnam, however, finds that the relationship between poverty and environmental risks is hard to generalize (Dasgupta *et al.*, 2005).

This paper revisits this relationship, taking advantage of more recent and higher-resolution data to examine four research questions: (RQ1) At the district-level is the incidence of poverty higher in high risk areas? (RQ2) At household-level do poor households face higher environmental risks? (RQ3) Does the relationship between household-level consumption and environmental risks vary between rural and urban areas? (RQ4) Do environmental risks relate to consumption differences between households and consumption changes over time? On their own, many of these questions have been addressed in the existing literature, but they have hardly been brought together to draw a comprehensive picture on the relationship between poverty and multiple environmental risks within a single country context using high-resolution data. We contribute to existing work by assessing how the relationship between poverty and multiple dimensions of environmental risk varies as a function of the channels through which poverty and risk interact, using a set of consistent data and methods while holding constant national context.

In this paper a number of recently-developed high-resolution datasets are used to separately assess eight environmental risks: (i) outdoor air pollution as a proxy for health risks through respiratory diseases (from Brauer *et al.*, 2016); (ii) tree cover loss as a proxy for the loss of forest resources and ecosystem functions (from Hansen *et al.*, 2013); (iii) land degradation as a proxy for productivity

[15] This number is based on the national poverty line, which is at is $3.49-a-day in 2011 PPP terms. By the $1.90-a-day poverty line, the extreme poverty rate was 3% in 2014.

decline and agricultural production risks (from Vu, Le and Vlek, 2014); (iv) slope as a proxy for areas vulnerable to erosion and landslides (from World Soil Database); (v-vi) long-term rainfall and temperature variability as a proxy for the variation of weather conditions (from the Climate Research Unit (CRU)); (vii) flood hazards as a proxy for rainfall- and coastal-related extreme events (from Bangalore, Smith and Veldkamp, 2019); and (viii) drought hazards as a proxy for drought events (from Winsemius *et al.*, 2018).

These data are combined with national maps showing the incidence of poverty at the district-level in 2010 (Lanjouw, Marra and Nguyen, 2013). In addition, this study computes environmental risks at the commune-level[16] to relate it to household information based on the latest Vietnam Household Living Standard Surveys (VHLSS) for 2010, 2012, and 2014 to examine whether environmental risks vary between households. Benefiting from the panel structure of these surveys, we also assess how environmental risks are related to consumption changes over time.

The remainder of this paper is structured as follows. Section 2.2. starts with deriving four research questions regarding the environmental risk and poverty linkages from existing literature. Section 2.3. describes the data used for the analysis of these research questions. The following four sections present the results. Section 2.4. shows the district-level overlay of environmental risks and poverty in 2010 (RQ1). Section 2.5. compares environmental risks at the commune-level across household groups (RQ2). Building on these analyses, Section 2.6. investigates differences in the relationship between household consumption levels and environmental risks between rural and urban areas (RQ3). Section 2.7 deepens these analyses by presenting results from various regression models explaining consumption differences between households and changes over time by environmental risks at the commune-level (RQ4). Section 2.8 concludes.

2.2. Conceptual framework

What do we know about the association between environmental risks and poverty? Intuition suggests that the two are positively associated: the higher the level of environmental risk, the higher the incidence of poverty (conversely, the lower the level of consumption). While empirical work lends some support to this claim, the relationship is multifaceted. In this section, we examine

[16] These are third-level administrative subdivisions, equivalent to a sub-district.

dimensions over which poverty and environmental risks might vary: (1) across risk, space, time, and scale, (2) between rural and urban areas, and (3) depending on the channels that determine the relationship. While the framework presented below does not fully address all the different dimensions and causes that underpin this complex relationship and is not supposed to integrate the full set of frameworks existing in this field, it does serve to motivate the empirical approach taken in this paper, which is to examine the multifaceted relationship across many risks in the single context of Vietnam.

2.2.1. Differences across risk, space, time, and scale

The literature associated with the Environmental Kuznets Curve (Panayotou, 1997) argues that the relationship between environmental risks and incomes might be inversely U-shaped, where environmental degradation increases with income levels up until a threshold is reached, after which degradation rates decrease. While consistent and rigorous evidence for this hypothesis has not been confirmed across a wide range of studies (for a review, see (Stern, 2004)), the framework highlights that the relationship between environmental risks and poverty is not monotonic or straightforward.

Nevertheless, a range of studies in the literature point at a positive relationship between environmental risks and poverty. At the global level, various studies show that environmental risks are higher in poorer countries (Barbier, 2010, 2015; Barrios, Bertinelli and Strobl, 2010; Sloan and Sayer, 2015). Studies at the sub-national level show that in some high-risk provinces and districts, the incidence of poverty is higher (Dasgupta *et al.*, 2005; Watmough *et al.*, 2016; Winsemius *et al.*, 2018). Similarly, at the household-level poorer households are found to have a higher exposure to some environmental risks than wealthier households (Ranger *et al.*, 2011; Akter and Mallick, 2013; Wodon *et al.*, 2014). Yet overall the findings are mixed and very context dependent.

First, the relationship depends on the specific risk being considered. For example, Hallegatte, Vogt-Schilb, *et al.* (2016) find that within countries, poorer areas are also more exposed to higher temperatures and drought, but not to floods. In terms of air pollution, poor people tend to be more exposed to both ambient (outdoor) but also indoor air pollution. For outdoor air pollution, studies focused mostly in the developed world find that polluting industry (including coal plants) often locate in lower-income neighborhoods (Braubach and Fairburn, 2010). In terms of indoor air pollution, poorer households in developing countries remain highly reliant on firewood or biomass for cooking, with women and children living in severe poverty having the highest exposure levels (Gordon *et al.*, 2014). One recent study that has looked at different environmental risks in Lao PDR finds strong connections with poverty for indoor air pollution and water quality, weak connections for deforestation and soil erosion and no connection for outdoor air pollution (Pasanen *et al.*, 2017).

For some environmental risks, even a negative relationship could exist resulting from economic development that has increased welfare levels, but degraded ecosystems, such as agricultural intensification or land expansion into natural areas. For example, rural households in sub-Saharan Africa, many of whom are poor, often rely on charcoal production to support growing demand in urban areas. While this contributes to poverty reduction through income-generation opportunities, it can also undermine soil and ecosystem stability and thereby agricultural production (Zulu and Richardson, 2013). Generally, deforestation can deprive rural households of natural resources they depend on, while at the same time providing new income sources (Chomitz *et al.*, 2007; Laurance, Sayer and Cassman, 2014). Where these contrasting effects coexist, it is possible that no significant relationship between degradation and poverty can be found.

More generally, within a country, the relationship between environmental risks and poverty is likely to change over time. For instance, Ebenstein *et al.* (2015) observe that over the period of 1991-2002, China grew incomes, reduced poverty, as well as increased air pollution. While high levels of air pollution increased mortality rates, economic growth and poverty reduction has also been associated with other health improvements that likely outweighed air pollution impacts. Also where ecosystems have already been degraded substantially in the past, further degradation will be limited. Such change over time may explain why areas with low soil erosion rates tend to be poorer (Naipal *et al.*, 2015).

Generally, poverty and environmental risks are unevenly distributed across space so that the relationship also depends on the places considered. The spatial distribution of environmental risks and poverty varies greatly between countries being shaped by a combination of agro-climatic and socio-economic conditions (Tucker *et al.*, 2015). Even in geographically similar countries, different patters can be observed, as shown by Dasgupta *et al.* (2005): While in Lao PDR there is a spatial correlation between poverty and all environmental risks, in Cambodia, poverty is only positively related to household-level risk factors, such as indoor air pollution and lack of access to adequate water and sanitation. For Vietnam, the authors only find a positive relationship with fragile soils and indoor air pollution.

Moreover, the scale of the analysis matters too. For example, Winsemius *et al.* (2018) find mixed evidence for a higher exposure of poor people to climate-related risks, like floods, when looking at poverty and risk levels at district-level. This finding could emerge as even within a district poverty and risks may be very unevenly distributed, so that more granular data at local-level from household surveys or case studies is needed. Indeed, looking at the same risk within a city – Mumbai, India, Ranger *et al.* (2011) find that poorer households in Mumbai are almost twice as exposed to floods.

However, analyses at higher resolution (e.g. household level) do not always show a positive association between environmental risks and poverty. Various studies show that there is no considerable difference between poor and wealthier households in their exposure to natural disaster risks (Del Ninno, 2001; Carter *et al.*, 2007; Opondo, 2013). These findings may be explained by the scale of the risk (in some cases, everyone is exposed as in the case of drought), but highlight that the relationship between poverty and exposure to environmental risks is context specific and depends on the type of hazard, local geography, institutions, and other mechanisms (Hallegatte, Vogt-Schilb, *et al.*, 2016).

So, overall the relationship between environmental risks and poverty remains an empirical question depending on the specific risk types and locations of interest and the scale considered. Accordingly, it needs to be tested whether patterns at provincial or district-level hold at the household-level. From this we derive our first and second research questions (RQ1 and RQ2) regarding the relationship between risks and poverty at different scales: (RQ1) Is the incidence of poverty (*P*) at the district-level higher in high risk areas than low risk areas? And (RQ2) at the household-level is the level of environmental risks (*R*) higher for poor than for non-poor households?

$Distrct - level$: $P_{High\ Risk\ Areas} > P_{Low\ Risk\ Areas}$ (RQ1)

$Household - level$: $R_{Poor} > R_{Non-Poor}$ (RQ2)

2.2.2. Differences between rural and urban areas

While various studies address the relationship between environmental risks and poverty in rural or urban areas, so far not many studies have examined differences between rural and urban spaces within the same country. Yet the extent of poverty as well as the spatial distribution of poor people in the countryside can differ substantially from cities, so that also the relationship between environmental risks and poverty could vary.

First of all, rural and urban areas differ in wealth levels and thus the incidence of poverty. Rural areas have generally lower consumption levels and poverty continues to be mostly a rural problem. In 2012, 78% of the global population in extreme poverty was living in rural areas (Olinto *et al.*, 2013). At the same time more and more destitute people move into cities in search for new opportunities so that inequalities in urban areas may be increasing (Fox, 2014). In many countries, rural populations are concentrated in fragile areas with higher levels of environmental risks, including land degradation (Barbier, 2010) and terrain ruggedness (Nunn and Puga, 2012). While some hotspots where high risks and poverty coincide can be found, environmental risks could also be higher in wealthier areas (Pasanen *et al.*, 2017). Many case studies show how marginal people in rural areas face environmental risks: for example, precipitation variability in the Peruvian Andes (Sietz, Mamani Choque and Lüdeke, 2012), floods in Senegal (Tschakert, 2007), drought in the Sahel (Sissoko *et al.*, 2011) and Northwest China (Li *et al.*, 2013), and cyclone-related saltwater intrusion in coastal Bangladesh (Rabbani, Rahman and Mainuddin, 2013). However, within rural areas, it is unclear whether poor people are more exposed than non-poor people. For instance, areas of higher forest cover loss might have lower poverty rates due to the income-generation opportunities associated with agricultural expansion or forest products (Chomitz *et al.*, 2007; Laurance, Sayer and Cassman, 2014).

In urban areas, the differences in exposure between poor and urban households could be more pronounced. Land scarcity is more pressing in urban areas, such that poor people (especially migrants) tend to locate in cheaper parts of the city and end up in slums (Fay, 2005). Often, these slum areas are characterized by higher environmental risks which are reflected in the price of land (Fay, 2005): for instance, Lall and Deichmann (2012) observe that parts of Bogota which experience the highest earthquake risk are also the cheapest, and where most of the poor locate. Accordingly, Winsemius *et al.* (2018) find a positive association between flood risk and poverty within urban areas, but do not find the same association at national-level, suggesting land prices to play a key role in determining exposure to environmental risks in cities.

More generally, differences across poverty levels in rural-urban settings may reflect the differing role of natural resources: whether they are primarily an input to production or to environmental quality (Panayotou, 2016). For instance, forests may be used as production input in rural areas so that they become increasingly degraded with higher wealth levels. In urban areas they may be seen as an environmental quality amenity and hence become more protected with higher wealth levels.

While these differing roles of natural resources, such as forests, can help explain findings in previous studies, such a framework cannot be used for all the environmental risks which we explore in this paper.

In this context, it is of interest to understand whether poor people live in risker places in rural and urban areas alike. Hence our third research question (RQ3) focuses on the difference in the relationship between environmental risks and poverty within rural and urban areas: Is the correlation between environmental risks (R) and poverty (P) positive (or negative for consumption levels) in rural as well as urban areas?

$$corr\ (R_{Rural}, P_{Rural}) > 0 \quad and\ corr\ (R_{Urban}, P_{Urban}) > 0 \quad \text{(RQ3)}$$

2.2.3. Different causes

So far, we have discussed whether there is a positive relationship between environmental risks and poverty, but not reasons behind this relationship. There can be multiple, interlinked channels that determine this relationship.

First, poor people and fragile areas could have shared vulnerabilities (Barrett, Travis and Dasgupta, 2011). For example, adverse hydroclimatic conditions could increase flood and drought hazards as well as ecosystem fragility, while also making livelihood activities, such as agriculture, more difficult, thereby resulting in lower incomes and consumption. For instance, dry climate zones in sub-Saharan Africa are vulnerable to precipitation deficits due to poor soils with low moisture storage capacity, and host rural people who remain poor as they are reliant on low-return and high-risk agricultural activities (Zimmerman and Carter, 2003; Barrios, Bertinelli and Strobl, 2010).

Second, poverty and environmental risks can be determined by common factors, such as institutional and market failures (Duraiappah, 1998; Barbier, 2010; Barrett, Travis and Dasgupta, 2011). For example, where property rights for land are missing, it is more likely that environmental resources, such as timber and fish, are overexploited (Baggio and Papyrakis, 2010). At the same time, powerful actors can oust poor people without land title from land leaving them without their main asset for wealth accumulation (Grainger and Costello, 2014). These institutional failures can also be related to market failures, where environmental services, such as water and soil regulation are not factored into market prices. While this lack of a price signal can lead to an underprovision of these services and hence higher risks, poor people managing ecosystems sustainably are not paid for this service provision.

Third, poverty could increase environmental risks (Duraiappah, 1998; Barbier, 2010; Barrett, Travis and Dasgupta, 2011). Many poor people – especially those in rural areas - depend on ecosystems for their livelihoods. For example, a systematic analysis of a 28-country data set shows that the in (sub-)tropical smallholder systems the poorest households derive more than half of their income from ecosystems and that this share is higher than for the wealthiest households (Angelsen *et al.*, 2014; Wunder, Noack and Angelsen, 2018). These incomes often play a role in consumption smoothing between seasons or as a coping mechanism when other incomes fail (Russell, Locatelli and Pramova, 2012). Although environmental extraction may not be a primary coping strategy (Wunder, Noack and Angelsen, 2018), ecosystem-based incomes are a substitute of other incomes and can stabilize total incomes when weather anomalies hit (Wunder, Noack and Angelsen, 2018). And where poor people lack other opportunities they may disproportionally resort to overexploiting natural resources (e.g., timber, fish or grassland) for short-term survival – a strategy that can cause poverty traps (Barbier, 2010; Barrett, Travis and Dasgupta, 2011).

Fourth, environmental risks could increase poverty (Duraiappah, 1998; Barbier, 2010; Barrett, Travis and Dasgupta, 2011). Where people depend on ecosystems as a safety net or for their base income, any environmental risks that undermine these functions will reduce their incomes and consumption. In (sub-)tropical smallholder systems an additional 13 percent of households would be in poverty without ecosystem-based incomes (Wunder, Noack and Angelsen, 2018). Moreover, the occurrence of natural disasters, can have a direct impact on welfare pushing people back into poverty or making it harder for poor people to escape poverty: for instance, in Bolivia, poverty rates increased by 12 percent in Trinidad city after the 2006 floods (Perez-De-Rada and Paz, 2008; Hallegatte, Vogt-Schilb, *et al.*, 2016). Households that lack ex-post coping mechanisms, often seek to mitigate risks ex-ante, for example through income diversification, thereby lowering average consumption and income (Bandyopadhyay and Skoufias, 2015). The extent to which the exposure to risks, without the materialization of an immediate direct impact on consumption and incomes, such as forest or land degradation can affect poverty has been less explored, and links are difficult to disentangle due to feedback loops and synergistic effects (Gerber, Nkonya and von Braun, 2014). Where the exposure to such risks has a negative effect on consumption and incomes, it is important to not only focus on risk response, but to derive strategies for risk reduction.

Hence it is important to understand whether risk exposure increases poverty or reduces consumption and income growth. Whereas the other channels of the environment-poverty nexus are not less important, our forth research question (RQ4) focuses on this channel:[17] Do environmental risks (R) increase poverty (P) (or reduce consumption)?

$P = f(R)\ or\ R \rightarrow P$ (RQ4)

2.3. Data

This study combines socioeconomic data from the Living Standard Measurement Surveys (VHLSS) in 2010, 2012, and 2014 measuring household-level consumption, and recent geo-spatial data measuring environmental risks at high resolution. These two data types are merged at the district and commune-levels.

2.3.1. Socioeconomic data

The VHLSS 2010, 2012, and 2014 are nationally and regionally representative and contain detailed information on individuals, households and communes.[18] In total 9,400 households nationwide are included in each round. A particularity of the VHLSS data is that half of these households were interviewed in two or three of these survey years. These households form a short-term '*Panel'* dataset to explain consumption changes over time. Other households were only observed in one of the three years. Using these cross-section data, treating all observations as independent observations provides a '*Pooled'* dataset to explain consumption differences between households. In rural areas 6,600-6,750 households from about 2,200 communes in each year are included ('Pooled' cross section), 1,400 of them are observed in all three rounds, 1,600 of them in 2010 and 2012, and 1,400 in 2012 and 2014 ('Panel' dataset). In urban areas the dataset covers 2,650-2,780 households in each year from 900 communes ('Pooled'), 500 of them are in all three rounds, 575 in 2010 and 2012, and 640 in 2012 and 2014 ('Panel').

The VHLSS data provide detailed information to estimate consumption values based on expenditure data. Section 2.4 uses district-level poverty maps. The ratio and number of people below the national poverty line[19] in each district is estimated using the consumption values from the VHLSS 2010 combined with the 15-percent sample of the 2009 Population and Housing Census as

[17] We also lack data to investigate the other three channels.
[18] These surveys are conducted by the General Statistics Office (GSO) with technical support from the World Bank in Vietnam.
[19] The national poverty line as calculated by GSO and the World Bank is used.

calculated by (Lanjouw, Marra and Nguyen, 2013). Sections 2.5 and 2.6 use consumption values from the VHLSS 2014 and Section 2.7 adds the data from the 2010 and 2012 surveys. All consumption values are calculated in line with the methodology for determining the national poverty line.[20]

This data from (Lanjouw, Marra and Nguyen, 2013) is used to examine RQ1. Through our own calculations, we use the 2014 VHLSS to examine RQ2 and RQ3, while we use all three rounds (2010, 2012, and 2014) to examine RQ4.

The household and commune surveys in the VHLSS 2010, 2012, and 2014 also include a wide array of data on socioeconomic conditions. Data at the individual-level include demographics, education, employment, health, and migration. At the household-level data comprise information on durables, assets, production, income, and participation in government programs. The commune surveys collect information about the commune characteristics including access to land, infrastructure and services. Based on these data a number of socioeconomic controls are constructed for the analyses based on household and commune surveys (the summary statistics can be found in Appendix B, Table 6).

2.3.2. Environmental risk data

Based on geo-spatial datasets, variables are constructed to measure environmental risk at district and commune-levels representing eight environmental risks. These variables are based on historical risk profiles measuring the area's exposure to fragile and severe conditions and not the actual environmental conditions at the time of the VHLSS surveys. The following variables are calculated to measure the area-weighted average of each environmental risk at the district and commune-level (the summary statistics can be found in Appendix B, Table 6): [21]

Air pollution is measured by the area-weighted mean of concentration (measured as micrograms per cubic meter) of particulate matter with a diameter of 2.5 micrometers or less (PM2.5) taking the 10 year-average value for 2000-2010. The data is based on satellite imagery using the total

[20] All consumption values are expressed in 2011 Purchasing Power Parity (PPP) values using data on the Consumer Price Index from the World Development Indicators.

[21] For the weather and temperature variability, due to the coarse resolution of the data, the value that the centroid of the commune is taken to represent these risks (rather than the area-weighted average).

column aerosol optical depth from the Moderate Resolution Imaging Spectroradiometer (MODIS) and Multiangle Imaging Spectroradiometer satellite instruments, which is combined with chemical transport model simulations, and ground measurements from 79 countries to produce a global spatial dataset with 0.1° × 0.1° resolution (Brauer *et al.*, 2016). PM2.5 includes dust, dirt, soot, smoke, and liquid droplets, which can lodge deeply into the lungs due to their small size. PM2.5 air pollution has been identified as a leading risk factor for global diseases (Forouzanfar *et al.*, 2016) as well as in Vietnam (Luong *et al.*, 2017).

Tree cover loss is calculated as the share of the area under tree cover in 2000 that suffered from a tree cover loss between 2000 and 2010. Tree cover is defined as canopy closure for all vegetation taller than 5m in height and is calculated from imagery from the Landsat 4, 5, 7, and 8 satellite data used to produce a global forest cover change map (Hansen *et al.*, 2013). Tree cover loss can be associated with a habitat disturbance and ecosystem service disruptions, which can make human landscapes more fragile to other impacts (Sodhi *et al.*, 2010). And it can also undermine the livelihoods of poor people highly dependent on forest timber and other resources (Angelsen *et al.*, 2014).

Land degradation is measured by the share of land area that experienced a significant biomass decline. This loss is calculated based on the inter-annual mean trend of Normalized Difference Vegetation Index based on data from Advanced Very High Resolution Radiometer (AVHRR) of the National Oceanic and Atmospheric Administration (NOAA) satellite between 1982 and 2006, which is corrected for climate effects to only measure human-induced degradation (Vu, Le and Vlek, 2014). Soil fertility, agricultural productivity and ultimately food security of smallholder farmers can be largely compromised on degraded lands (von Braun *et al.*, 2013).

Slope is measured by the area-weighted average of slope categories. This variable is calculated based on data from the Harmonized World Soil Database version 1.2 with eight slope classes: 1 for least steep (elevation of 0-0.5 percent) and 8 for most steep slope (elevation greater than 45 percent).[22] Steep slopes are much more prone to surface water runoff and soil erosion - particularly in areas affected by heavy rainfalls and tree cover loss (Sidle *et al.*, 2006; Vezina, Bonn and Van,

[22] http://www.fao.org/soils-portal/soil-survey/soil-maps-and-databases/harmonized-world-soil-database-v12/en/

2006). At the same time tropical cyclones can cause fatal landslides, which already result in significant loss of life in mountainous areas of South and East Asia (Petley, 2010, 2012).

Rainfall variability is defined as the 1981-2010 standard deviation of monthly rainfall levels measured in millimeters (mm). The variable is constructed from the global CRU TS3.21 dataset from the University of East Anglia, containing a long-term time series of monthly rainfall levels at 0.5x0.5 grid resolution, which was produced using statistical interpolation based on data from 4,000 weather stations (Harris *et al.*, 2014). Higher historic rainfall variability indicates higher inter-annual and intra-annual variation of precipitation and a higher incidence of extremely wet and dry conditions. In Vietnam excessive rainfall can have negative income effects for poor households, while wealthier households are more negatively affected by the lack of rainfall (Narloch, 2016).

Temperature variability is measured by the standard deviation of mean annual temperatures across the time period 1981-2010 in Celsius degree (C). The underlying data also comes from the global CRU TS3.21 dataset including a long-term time series of monthly mean temperature values at 0.5x0.5 grid resolution. Temperature variability indicates varying temperature conditions between seasons and years and a higher incidence of heat waves and cold spells. In Vietnam hot conditions have generally an income-reducing effect for rural households (Narloch, 2016).

Flood hazards are represented by the share of area at risk of a flood event (inundation depth > 0) with a 25 year return period under historical conditions.[23] The measures are based on the inundation depth estimated by state-of-the art flood models at a grid cell level of 3 arc-seconds combining coastal surge hazard layers, along with pluvial and fluvial layers (Bangalore, Smith and Veldkamp, 2019). In Vietnam flood events have been shown to significantly reduce welfare and increase poverty (Thomas *et al.*, 2010; Bui *et al.*, 2014; Arouri, Nguyen and Youssef, 2015).

Drought hazards are measured by the area-weighted intensity of drought conditions. This intensity is a measure of hydrological droughts and expressed by the number of months of long-term mean discharge which would be needed to overcome the maximum accumulated deficit volume during dry months (Winsemius *et al.*, 2018). Drought events in Vietnam are associated with agricultural

[23] A 25 year return period corresponds to 0.04 annual probability of occurrence.

production losses, negative welfare impacts and poverty (Thomas *et al.*, 2010; Bui *et al.*, 2014; Arouri, Nguyen and Youssef, 2015).

2.3.3. Data limitations

There are several limitations in using these datasets for the purpose of this study. Nevertheless, these data allow to provide some interesting insights into our research questions.

First, while the data is representative for socio-economic conditions of the entire country, its eight regions and rural and urban areas, it is not for the different environmental risk profiles across communes. Building on the living standard measurement survey (LSMS) methodology, the sampling strategy follows a three-stage stratified cluster design, whereby about a third of all communes are selected and divided into three areas from which one area is chosen and three households are interviewed from these areas (ISM; SINFONICA, 2015). Due to this strategy, not the entire variety of communes is represented and communes in very remote areas and with extreme risk profiles, for example those in remote mountainous areas, natural forests or islands, may not be captured at all or underrepresented. This sampling bias could lead to an underrepresentation of very high risks areas

Second, all the environmental risks variables on hand are time-invariant (i.e. have the same value for all the years with household survey data) as they are based on historic risk profiles. On the one hand, it would be preferable to measure actual conditions during the survey years to control for any changes between the years, which is not possible as such data was not available when this study was prepared. On the other hand, measuring environmental risks based on past conditions minimizes causality problems, whereby consumption and income can determine current environmental conditions. However, the data cannot address any omitted variables bias, which leads to some endogeneity concerns (as will be discussed in Section 2.5.).

Third, most of the environmental risks variables are based on global datasets using global models or data, which are not necessarily representative for the specific conditions within the districts and communes in Vietnam. Optimally such variables would be measured based on ground station data, which however is not readily available. Although these are important limitations that require further work, the available data can provide some first insights into the relationships between environmental risks and poverty.

2.4. Incidence of poverty in low and high risk districts

The following spatial analysis addresses the first research question (*RQ1*) exploring whether the incidence of poverty is higher in high risk districts than in low risk districts. Although the exposure of poor people to environmental conditions depends on their exact location in the district, such a district-level analysis helps to identify spatial hotspots where poverty coincides with high environmental risk levels.

2.4.1. Methods

National maps are produced that show the extent of environmental risks and poverty at district-level. Based on the 2010 poverty rate maps from (Lanjouw, Marra and Nguyen, 2013), all districts are classified into high, medium, or low poverty categories with an equal number of districts in each group using the poverty rate (i.e. relative poverty) and number of poor people (i.e. absolute poverty) in each district. Similarly, all environmental risk variables are separately calculated at district-level to categorize districts into high, medium and low risk districts.[24]

We chose this relative categorization (three terciles for each environmental risk, with equal number of districts in each category) as guidance on creating risk-specific categories was largely unavailable. While thresholds exist for air pollution based on guidance from the WHO which allow categorization into low, medium, and high (Brauer *et al.*, 2016), the other 7 environmental risks do not have clear guidelines on threshold values. Rather than create our own threshold values, we decided on a relative approach to split the districts evenly into three terciles for each risk, consistent with the poverty categorization. We are unaware of published studies creating similar "bivariate choropleth" maps in this field, although a recent paper on poverty and access to healthcare follows the same categorization rule that we choose (Tansley *et al.*, 2017). Given our RQ1 aims to examine the confluence of poverty and risk (e.g. high poverty and high risk districts), we argue that this categorization is suitable. The environmental risk maps are then overlaid with the relative (Figure 4, Panel a) and absolute poverty maps (Figure 4, Panel b). In addition, the poverty rate and number of poor people in each risk category is calculated (Table 12 and Appendix B, Table 7).

---

[24] For land degradation, due to the large number of districts experiencing no loss (0), the low risk category includes all districts with 0 loss (243 districts), and the rest of the districts are split into the medium and high categories.

Table 12. Poverty rates in 2010 across environmental risk categories at district-level

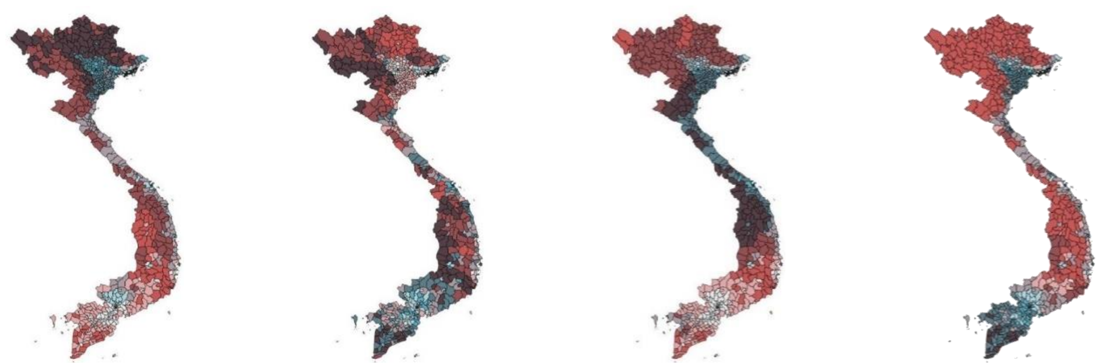
Risk category	Air pollution	Tree cover loss	Land degradation	Slope	Rainfall Variability	Temperature Variability	Flood hazards	Drought hazards
Low	21.3%	13.5%	19.5%	14.6%	20.6%	18.0%	49.1%	21.4%
Medium	33.2%	29.1%	32.6%	16.8%	35.6%	32.2%	16.7%	28.3%
High	26.9%	38.6%	30.0%	49.7%	25.5%	31.5%	15.4%	31.6%
ANOVA								
F	16.83	92.14	24.11	361.41	28.49	31.56	322.54	12.71
Prob>F	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

Notes: Table shows the poverty headcount rate across the three environmental risks categories as calculated for Figure 4 and statistics from a one-way analysis-of-variance (ANOVA), which assess whether the difference in poverty rates across risk categories is statistically significant.

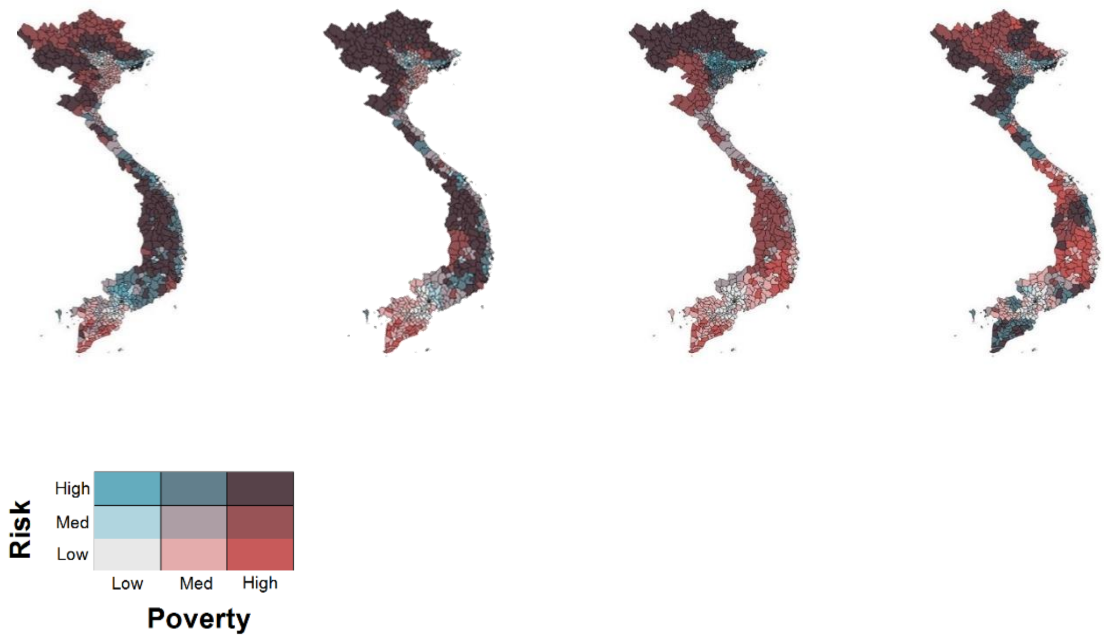
Figure 4. Overlay of poverty in 2010 and environmental risk categories at district-level

a. Poverty rates

a. Air pollution **c. Land degradation** **e. Rainfall variability** **g. Flood**

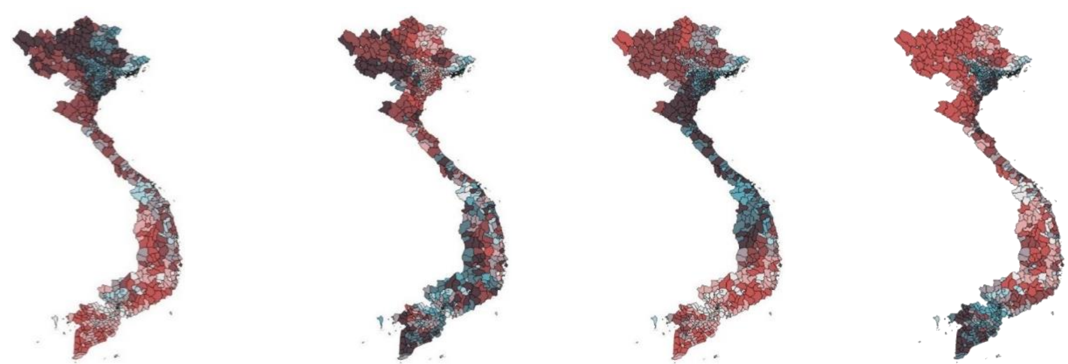


b. Tree cover loss **d. Slope** **f. Temperature variability** **h. Drought**

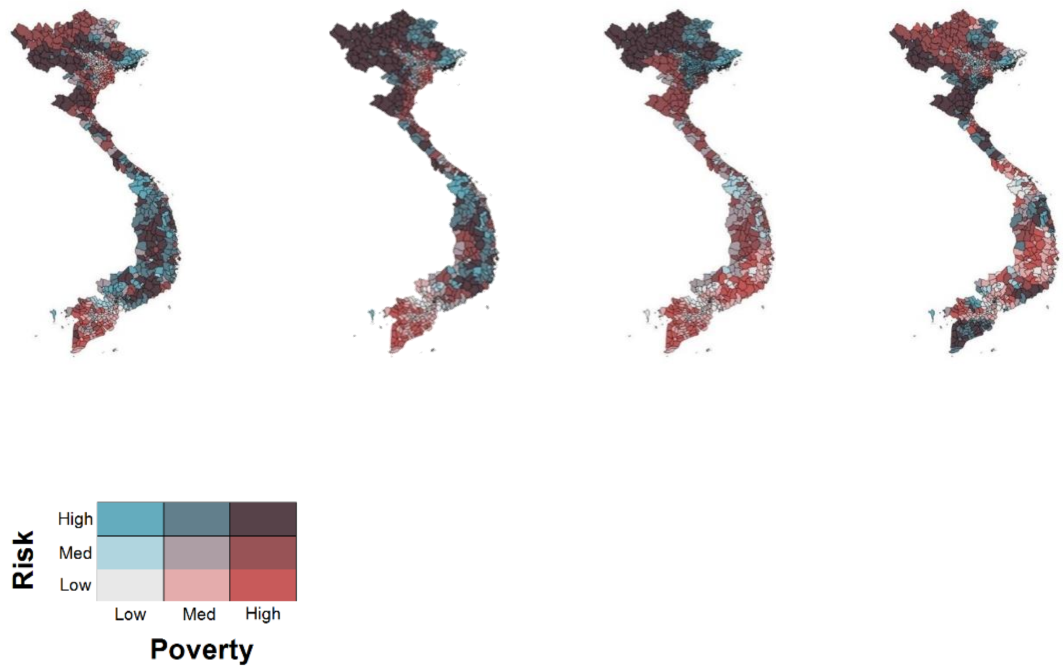


b. Number of poor people

a. Air pollution **c. Land degradation** **e. Rainfall variability** **g. Flood**



b. Tree cover loss **d. Slope** **f. Temperature variability** **h. Drought**



Notes: Figures the district-level overlay of poverty and environmental risks maps. Panel a shows the rate of people in each district below the national poverty line with Poverty- Low = <14%, Medium = 14-28%, High = >28%. Panel b shows the number of people in each district below the national poverty line with Poverty - Low = <15,000, Medium = 15,000-29,000, High = >29,000. Districts are classified based on the following categories for environmental risks (calculated as three terciles, with equal number of districts in each category): Air pollution – Low = < 9.11 , Medium = 9.11-23.75 High = >23.75 based on area-weighted PM2.5 pollution levels (micrograms per cubic meter); Tree cover loss - < 0.21 , Medium = 0.21-2.28 High = >2.28 based on share of forest area affected by tree cover loss; Land degradation – Low = 0, Medium = 0.01-20.14 High = >20.14 based on area share affected by biomass loss; Slope - Low = < 2.7, Medium = 2.7-5.3 High = >5.3 based on area-weighted average of slope category; Rainfall variability - < 45.50, Medium = 45.50-61.44 High = >61.44 based on the long-term standard deviation of monthly rainfall; Temperature variability – < 0.61, Medium = 0.61-0.85 High = >0.85 based on based on the long-term standard deviation of monthly mean temperature; Flood hazards: Low = <10%, Medium = 10-30%, High = 30%> based on the area shared at risk of a 25-year return period flood; Drought hazards: - < 0.87, Medium = 0.87-1.04 High = >1.04 based on number of months to overcome water deficit during dry months.

2.4.2. Results

Across Vietnam there is considerable spatial variation in poverty and environmental risks and there are some large hotspots of high environmental risks and poverty. As indicated by the dark districts in the maps (Figure 4, Panel a), Northern districts face a combination of high poverty and high air pollution, tree cover loss, steep slopes, temperature variability and drought hazards. In the Central Highlands, tree cover loss, steep slopes and rainfall variability are higher in poorer districts. And in the Mekong River Delta there are a few poorer districts that face high land degradation, flood and drought hazards. When looking at absolute poverty numbers, a few shifts in these patterns emerge from sparsely populated districts, as for example, in the Central Highlands, to districts with larger concentrations of poor people (Figure 4, Panel b).

Generally, in high risk districts the poverty rate is higher than in low risk districts. The difference is most pronounced for tree cover loss and slope, but also significant for all the other environmental risks (Table 12). Only for flood hazards, poverty is significantly higher in low risk districts than in high risk districts. This observation can be explained with the higher incidence of flood hazards in prosperous coastal regions and river deltas (Figure 4, also see (Bangalore, Smith and Veldkamp, 2019).

Similarly, a high number of people below the poverty line are concentrated in high risk areas. And the number of poor people is significantly higher in high risk than in low risk areas, with an average of about 30,000 people living in districts with high tree cover loss and land degradation, steep slopes and high drought hazards (compared to 20,000 people in low risk districts) (Appendix B, Table 7). And even an average of 20,000 poor people live in districts with high flood hazards, implying that floods can still affect a high number of poor people, even though they affect relatively wealthier districts.

2.5. Risks among different household groups

This section addresses the second research question (*RQ2*) exploring whether environmental risks at the commune-level are significantly different between household groups. Although there can be considerable variation in environmental conditions within communes, environmental risks measured at commune-level can measure the wider risk environment households are exposed to.[25]

2.5.1. Methods

To compare environmental risks across households, the level of environmental risks is calculated for each commune. Whereas the geographical location of communes is known, the household data from the VHLSS is not geo-coded so that it is not possible to track the exact location of households within communes. In addition to the commune-level absolute risk value, the standardized value of

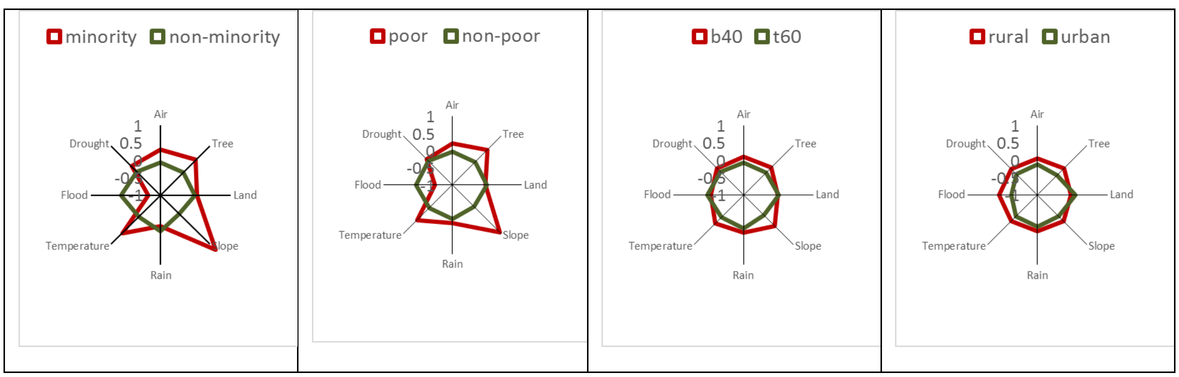
[25] Furthermore, as communes are about a fifth of the size of districts, there is likely to be considerably less variation in environmental risk within communes than within districts.

each environmental risk is calculated.[26] Such normalization produces a scaled version of the original value which allows comparison between the different risk variables. The absolute (Appendix B, Table 8) and standardized values (Figure 5) are then compared across household groups in 2014, distinguishing between (i) households from ethnic minorities vs. non-minorities, (ii) households below ('poor') and above the poverty line ('non-poor'), (iii) households in the bottom two consumption quintiles ('b40') and the top three consumption quintiles ('t60') and (iv) households in rural ('rural') and urban ('urban') areas.

2.5.2. Results

There are considerable differences in risk profiles between different groups. Households from ethnic minorities, those below the poverty line and those in the lowest two consumption quintiles live in much riskier communes (Figure 5). It can be seen that the differences are most pronounced for the minority versus non-minority households. Poor households have a similar risk profile, also because more than 50 percent of poor households belong to ethnic minorities. Accordingly, households that face marginalization due to poverty or ethnicity are also disproportionally exposed to multiple, possibly interlinked environmental risks.

Figure 5. Environmental risk profiles across socio-economic groups in 2014



Notes: Spider diagrams show the mean risk values of households in each socio-economic group: households that belong to ethnic minorities (Minority), households below the national poverty line (Poor), the households in the lower two consumption quintiles (Bottom40), and households in the upper three consumption quintiles (Top60). Environmental risk values are standardized by subtracting their mean and

[26] From each commune value the mean is subtracted and divided by the standard deviation, producing a value distribution with the mean of 0 and a standard deviation of 1.

dividing by the standard deviation. Descriptive statistics of all environmental risks variables with original units are shown in Appendix B, Table 6.

For most environmental risks the differences between groups are significant. Ethnic minorities, poor households and those in the lowest two consumption quintiles face much higher air pollution, tree cover loss, steeper slopes, rainfall and temperature variability and drought hazards than their wealthier counterparts (Appendix B, Table 8). They are, however, exposed to lower flood hazards confirming results from Section 2.3 that flood risks are higher in more prosperous districts. These risk levels are very similar among groups in 2010, 2012 and 2014 suggesting that movement out of high risk zones has remained very limited over time (Narloch and Bangalore, 2016).

Looking at differences between rural and urban households, aside from land degradation all other risk levels are much higher (Appendix B, Table 8). Accordingly, the general risk profile is also more extreme for rural than urban households (Figure 5). This finding is expected as most of the ethnic minorities, poor households and those in the lowest consumption quintiles are rural households. Yet this finding does not imply that within rural or urban areas poorer households are also more exposed than their wealthier counterparts, which will be further analyzed in the next section.

2.6. Risks and poverty within rural and urban areas

This section addresses the third research question (*RQ3*) touching upon the last section to show whether the relationship between environmental risks at the commune-level and household-level consumption levels are different between rural and urban areas.

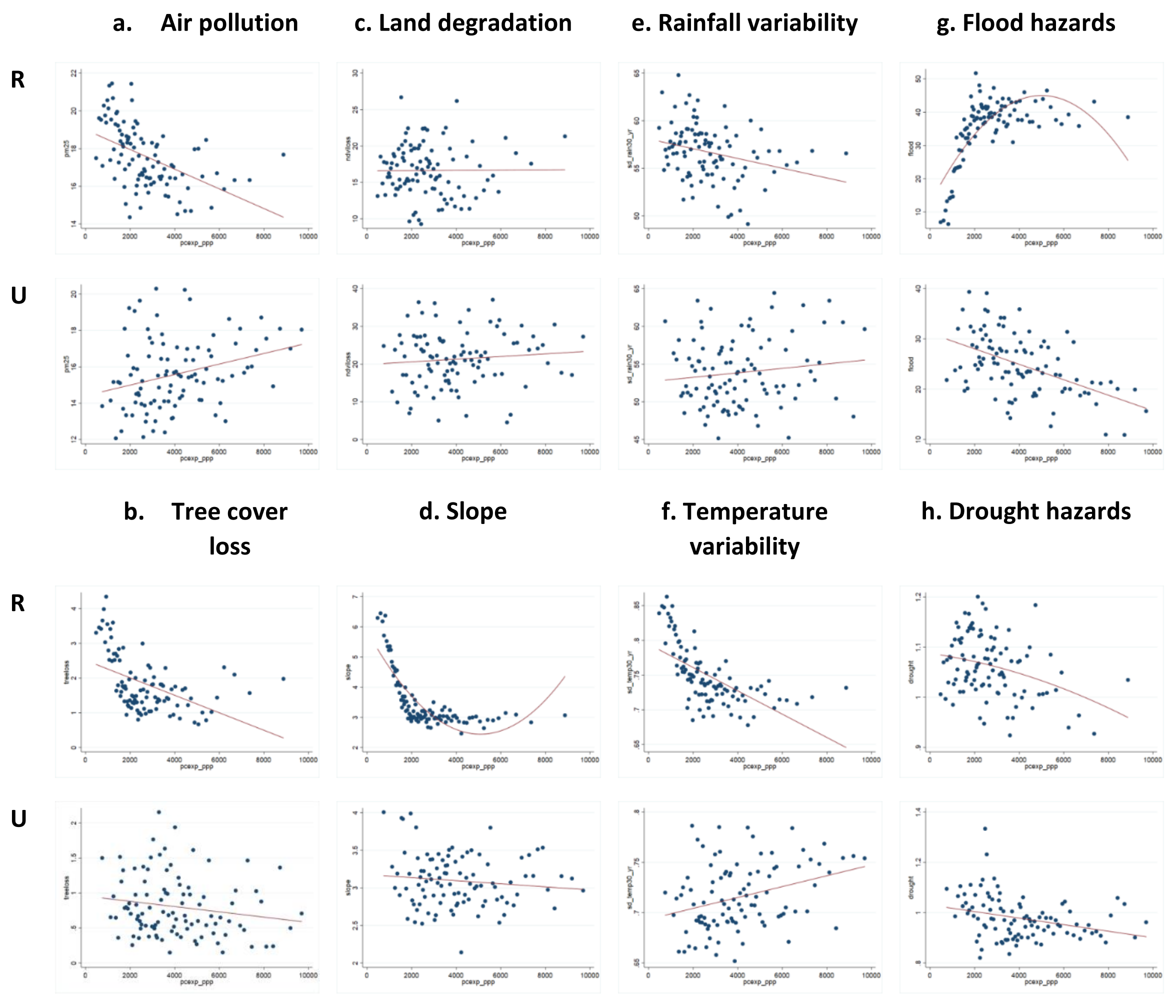
2.6.1. Methods

The data is split into a rural and urban subsample to calculate correlation coefficients and run nonparametric analyses. For both sub-samples the correlation coefficient between households-level consumption based on the expenditure data from the VHLSS 2014 and commune-level risks variables is calculated (Appendix B, Table 9). Households are further divided into consumption percentiles (100 groups) to show the mean level of consumption (x-axis) and of environmental risks (y-axis) for each percentile for the rural and urban sub-sample (Figure 6).

2.6.2. Results

In both rural and urban areas, poorer households tend to live in communes with higher environmental risks. While for most environmental risks the correlation with per-capita consumption is negative within rural as well as urban areas, the degree of correlation varies between rural and urban areas (Appendix B, Table 9). In rural areas households in the lower percentiles are more exposed to air pollution, tree cover loss, steep slopes, rainfall and temperature variability, but less exposed to flood hazards (Figure 6). In urban areas, poorer households are more concentrated in areas with high tree cover loss, slope, flood and drought hazards, while living in places with lower air pollution and temperature variability (Figure 6).

Figure 6. Environmental risks across consumption percentiles within rural (R) and urban (U) zones in 2014



Notes: Binned scatterplots show the non-parametric relationship between consumption and environmental risks with each data point representing a consumption percentile. X-axis shows the mean level of per-capita expenditure in 2014 (e.g. measure of poverty) for each percentile. Y-axis indicates the mean level of environmental risks for each percentile. R denotes Rural and U denotes Urban.

Some remarkable differences emerge between rural and urban areas (Figure 6 and Appendix B, Table 9). Poorer households seem to be less exposed to flood hazards in rural areas, but more so in urban areas. This finding could reflect that many of the rural poor live in the mountainous areas where flood hazards are generally lower than in low-lying, but wealthier river deltas and coastal zones, while many of the urban poor are pushed into high risk areas due to land constraints in the cities. This finding is consistent with a recent global analysis of flood exposure and poverty in 52 countries, which reports ambivalent results for flood exposure at the country level, but shows a strong signal of over-exposure of the poor when only focusing on urban areas (Winsemius *et al.*, 2015). Also poorer households are more exposed to air pollution and temperature variability in rural areas, but less in urban areas. Possibly, wealthier households live in rapidly developing urban areas with heavy traffic and industrial congestion, thereby being more exposed to air pollution.

2.7. Environmental risks and consumption differences and changes

To address the forth research question (*RQ4*) this section investigates how environmental risks relate to consumption differences between households and consumption changes over time, applying a set of regression models. Notwithstanding several limitations, these analyses help provide first insights into the relationship between environmental risks and poverty.

2.7.1. Methods

Two sets of regression models are fitted to estimate the effect of environmental risks on consumption differences between households in the 'Pooled' cross-section and on consumption changes over time for the 'Panel' dataset. Using Ordinary Least Squares (OLS) estimators the following equation is fitted in the *Pooled* and *Panel* models:

Equation 3. Regression model to estimate the effect of environmental risks on consumption

$$Y_{i_{j_t}} = \alpha + \beta R_j + \delta Z_j + \omega W_{j_t} + \gamma X_{i_{j_t}} + \lambda T_t + u_{i_{j_t}}$$

where Y denotes per-capita consumption observed for household *i* in commune *j* in year *t (i.e.* 2010, 2012, 2014*)* in the *Pooled* model and the change in per-capita consumption for household *i* between years *t* (*i.e.* 2010-2012 and 2012-2014) in the *Panel* model. *R* measures the environmental risk profile in commune *j* , which is time-invariant, as described in Section 2.3. *β* is the parameter of interest that indicates the consumption effect of environmental risks. *Z* includes commune-specific geo-spatial controls that do not change over time, such as proximity to cities and roads and the long-term average of rainfall and temperature. *W* includes a set of time-variant commune variables measured by rainfall and temperature levels in the survey years 2010, 2012 and 2014 for the *Pooled* model and by the change for 2010-12 and 2012-14 for the *Panel* model. *X* is a set of household-specific controls that vary over time, such the levels of education, labor and land endowments in survey years for the *Pooled* model and the difference between years for the *Panel* model. *X* also includes time-invariant demographic characteristics, such ethnicity and gender.[27] *T* measures time-fixed effects to neutralize common trends over time. *U* is a random, idiosyncratic error term. The descriptive statistics of these variables are provided in Appendix B, Table 6.

A main concern when fitting models to estimate weather impacts on economic outcomes is endogeneity bias. Problems of reverse causation (e.g. high consumption growth in one place leading to environmental depletion) are minimized by taking the historic risk profile and not measuring actual environmental conditions in the survey years. Yet the model is likely to suffer from omitted variable bias caused by the potential correlation of risk variables with other commune characteristics that determine living standards.[28] To minimize omitted variable bias in this study, a set of observable commune characteristics is included that are likely to influence risk profiles and living standards, such as average rainfall and temperature conditions and distance from cities and roads. Nonetheless, to the extent that other, non-observable commune characteristics also determine risks and consumption, the estimates of β are biased. For instance, some communes may differ on the level of trust, which is likely to be associated with higher living standards and better environmental outcomes. In this case, our estimates would be biased upwards.

These regression models are estimated for various sub-samples to disentangle varying effects between groups, zones and years. For both the *Pooled* and *Pane*l models, we first estimate the consumption effect one by one and then include all risk variables altogether. While only the results of the latter are presented, most of the results hold in the models with only one risk variable. Robust standard errors are estimated by clustering at the commune-level in order to account for spatial correlation. For the *Pooled* model the natural logarithm of per-capita consumption is used in the regression to normalize the skewed distribution of consumption (i.e. many observations of

[27] The datasets allow to control for land and labor inputs of some activities. These are, however, not included in the regression analyses due to potential endogeneity biases, as higher risks also determine land and labor inputs.

[28] This omitted variable bias could be minimized by fitting a fixed-effects linear model using a within-regression estimator based on the Panel dataset as in Narloch (2016). Yet such a model does not allow to disentangle the impact of time-invariant risk factors, such as the environmental risk variables so that it cannot be applied for the purpose of this study.

low consumption levels and a few observations of very high consumption levels).[29] This skewed distribution is common in household surveys in developing countries (Thomas *et al.*, 2010; Heltberg, Oviedo and Talukdar, 2015; Fraval *et al.*, 2019), including in Vietnam (ISM; SINFONICA, 2015), with the literature and guidance reports suggesting log transformations of consumption to aid interpretability.

2.7.2. Results

Environmental risks are related to consumption differences between households. The results for environmental risks are presented in Table 13, with the full results (including for all controls) found in Appendix B, Table 10. Across all households from the Pooled cross-section, those households, who live in communes with steeper slope, higher rainfall and temperature variability and flood and drought hazards have significantly lower consumption (Table 13 – Column All). Only for land degradation there seems to be a positive correlation, which may indicate communes of intensive agricultural expansion and growth, which could be positively associated to wealth in the short-term. This finding points at the possible estimation biases from unobservable factors, which do not allow to establish clear causal relationships with these data, especially as both time-varying and time-invariant variables are employed at multiple scales.

Table 13. The effect of environmental risks on consumption differences between 'Pooled' households in 2010, 2012 and 2014

Dependent variable: Ln of per-capita expenditure	All	Poor	B40	Rural	Rural 2014	Urban	Urban 2014
Air pollution	0.000511	0.00176	**0.00255****	**0.00425*****	0.00220	-0.00178	-0.00315
	(0.32)	(1.01)	(2.18)	(2.60)	(1.14)	(-0.55)	(-0.81)
Tree cover loss	-0.00179	0.00229	-0.00125	0.00132	-0.00248	-0.00705	-0.00500
	(-1.18)	(1.60)	(-1.10)	(0.88)	(-1.37)	(-1.35)	(-0.74)
Land degradation	**0.000300***	-0.000256	0.000171	0.000108	0.000197	-0.0000844	-0.00000728
	(1.87)	(-1.27)	(1.26)	(0.55)	(0.81)	(-0.33)	(-0.02)
Slope	**-0.0213*****	**-0.0148****	**-0.0102****	**-0.0101***	-0.00389	-0.0164	-0.0117
	(-3.98)	(-2.32)	(-2.43)	(-1.71)	(-0.55)	(-1.58)	(-0.88)

[29] Taking the natural logarithm also improves the explanatory power of the various models and brings ease in interpreting the coefficients as percentage change of the outcome variable.

Rainfall variability	-0.00276***	-0.00120	-0.000420	-0.00281***	-0.00499***	-0.000672	-0.000577
	(-4.48)	(-1.46)	(-0.75)	(-4.21)	(-4.73)	(-0.58)	(-0.33)
Temperature variability	-0.265*	0.0382	-0.212*	-0.454***	-0.0311	0.113	0.318
	(-1.88)	(0.22)	(-1.68)	(-2.98)	(-0.17)	(0.42)	(0.95)
Flood hazard	-0.00144***	-0.000202	0.00000600	0.000183	0.000441	-0.00184***	-0.00201***
	(-5.97)	(-0.74)	(0.04)	(0.67)	(1.43)	(-4.05)	(-3.96)
Drought hazard	-0.0335***	0.00353	-0.0199**	-0.0428***	-0.0379**	-0.0380	-0.0477
	(-2.67)	(0.26)	(-2.01)	(-3.17)	(-2.29)	(-1.59)	(-1.56)
Controls (see Appendix B, Table 10)	yes	yes	yes	yes	yes	yes	yes
_cons	8.024***	6.487***	6.849***	7.513***	6.677***	7.580***	7.225***
	(35.15)	(25.99)	(32.66)	(30.26)	(20.11)	(18.23)	(13.23)
N	27698	3811	11025	19748	6492	7950	2715
R-sq	0.432	0.204	0.365	0.399	0.409	0.316	0.305

Notes: Table indicates coefficients estimated from 'Pooled' Cross-section model using Ordinary Least Squares. * 0.10 ** 0.05 *** 0.01 significance level. Values in parentheses indicate standard errors corrected for cluster correlation at commune-level. B40 denotes households in the bottom two consumption quintiles. Controls include current rainfall, current temperature, long-term rainfall, long-term temperature, distance city, distance road, area agriculture, area forest, area water surface, workforce, education, age head, female head, minority, year 2012, and year 2014. Full results (including for controls) presented in Appendix B, Table 10.

Within the various groups different risk factors matter for consumption levels. As differences in per-capita expenditure between households below the poverty line are modest, it is not surprising that generally fewer variables are significant for the sub-sample of poor households and that the models have a much lower explanatory power (Table 13 – Poor). The role of environmental risks in explaining consumption differences for households in the lowest two consumption quintiles follow the overall pattern, but there is a significant positive relationship with air pollution (Table 13 – B40). The same finding emerges from the rural subsample (Table 13 – Rural). Possibly rural households in communes closer to urban centers are wealthier, while being more exposed to air pollution than their counterparts in remote rural areas. Whereas in rural areas slope, rainfall and temperature variability and droughts are negatively related to consumption, in urban areas floods have a significant negative effect (Table 13 – Urban) confirming results in Section 2.6. Very similar findings appear for the 2014 subsample, indicating that there are no considerable differences between the three survey years (Table 13 – Urban 2014 and Rural 2014).[30]

Changes in consumption over time are only related to a limited number of risk variables and some show unexpected signs. The results for environmental risks are presented in Table 14, with the full results (including for all controls) found in Appendix B, Table 11. Temperature variability relates to a significantly higher consumption growth mainly driven by urban households between 2012 and 2014 (Table 14 – All and Urban 2012-14), while for households within the lowest two consumption quintiles, floods seem to have a positive effect on consumption growth (Table 14 – B40). These effects can capture favorable weather conditions between 2010 and 2014 in these risk prone areas. For example, as long as no flood or extreme heat happens, people living in flood plains or hotter zones can benefit from more profitable activities, such as floating rice or coffee cultivation. Higher land degradation is positively related to consumption growth in the rural subsample (Table 14 – Rural). This effect may capture unobserved factors that facilitate higher consumption growth, such as non-farm income opportunities in more degraded areas.

Table 14. The effect of environmental risks on consumption changes over time of 'Panel' households in 2010-12 and 2012-14

	Dependent variable: Change in per-capita expenditure						
	All	Poor	B40	Rural	Rural 2012-14	Urban	Urban 2012-14
Air pollution	**-12.37***	6.040	-0.0273	-4.189	-3.224	-34.10	-41.43
	(-1.91)	(1.52)	(-0.01)	(-0.86)	(-0.42)	(-1.58)	(-1.61)
Tree cover loss	-4.166	-1.376	-5.865	-4.605	**-13.16***	-0.818	-5.808
	(-0.73)	(-0.40)	(-1.63)	(-0.83)	(-1.92)	(-0.03)	(-0.11)
Land degradation	0.262	0.554	0.0442	**1.555****	0.542	-2.124	-2.237
	(0.33)	(0.96)	(0.09)	(2.13)	(0.47)	(-1.15)	(-0.72)
Slope	8.262	-17.99	4.052	-2.692	-1.030	38.90	67.10
	(0.43)	(-0.93)	(0.30)	(-0.16)	(-0.04)	(0.73)	(0.91)
Rainfall variability	-0.162	**-3.139***	0.445	-0.349	-0.132	-0.250	**-15.18***
	(-0.08)	(-1.71)	(0.32)	(-0.17)	(-0.04)	(-0.04)	(-1.72)
Temperature variability	**974.5***	193.5	182.0	377.8	403.0	2582.1	**6247.0****
	(1.74)	(0.61)	(0.67)	(0.88)	(0.63)	(1.42)	(2.40)
Flood hazard	-0.390	0.110	**1.533****	-0.0430	-0.602	-4.244	-3.479
	(-0.38)	(0.10)	(2.33)	(-0.05)	(-0.45)	(-1.30)	(-0.99)

[30] These results also hold when including the risk variables one-by-one and not the whole set.

Drought hazard	-2.442	**-121.1****	-64.84	-35.65	53.63	137.7	58.11
	(-0.05)	(-2.06)	(-1.55)	(-0.72)	(0.67)	(0.89)	(0.23)
Control variables (results in Appendix B, Table 11)	yes	yes	yes	yes	yes	yes	yes
_cons	-1943.4**	188.1	-784.5*	-542.3	-1461.6	-5280.7**	-12451.3***
	(-2.25)	(0.39)	(-1.70)	(-0.73)	(-1.38)	(-2.17)	(-3.05)
N	7957	952	3250	5804	2792	2153	1109
R-sq	0.017	0.171	0.076	0.020	0.035	0.029	0.050

Notes: Table indicates coefficients estimated from 'Panel' model using Ordinary Least Squares and differences over time for time-variant variables. * 0.10 ** 0.05 *** 0.01 significance level. Values in parentheses indicate standard errors corrected for cluster correlation at commune-level. B40 denotes households in the bottom two consumption quintiles. Controls include current rainfall, current temperature, long-term rainfall, long-term temperature, distance city, distance road, area agriculture, area forest, area water surface, workforce, education, age head, female head, minority, year 2012, and year 2014. Full results (including for controls) presented in Appendix B, Table 11)

Yet a number of risk variables is related to lower consumption growth between 2010 and 2014 indicating a negative effect. Households in communes with higher PM2.5 concentration levels have a significantly lower consumption growth (Table 14 – All). Lower consumption growth in 2012-14 is related to tree cover loss in rural zones and to rainfall variability in urban zones (Table 14 – Rural 2012-14 and Urban 2012-14). Among poor households those living in communes with higher rainfall variability and drought hazards have a significantly lower consumption growth (Table 14 – Poor), which is an alarming result suggesting that increasing rainfall variability and recurring droughts mainly have negative impacts on those that are already in a destitute situation.

Some interesting findings also appear from the results for the control variables (full results presented in Appendix B Tables 10 and 11). Higher levels of current rainfall and lower current temperatures are related to higher consumption levels and growth over time, which is broadly in line with the results in Narloch (2016).[31] While larger distance to cities and roads relate to lower consumption levels as would be expected, it is related to higher consumption growth over time – possibly indicating a catching-up of more remote communes. Having access to larger agricultural and water surface areas has a positive effect on both consumption levels and growth over time. Ethnic minorities have lower consumption levels and also experience lower consumption growth

[31] Some of the results in Narloch (2016) differ as the analyses reduce omitted variable biases through fixed effects regressions.

when compared to wealthier households. However, within the sub-samples of households below the poverty line and in the lowest two consumption quintiles, ethnic minorities have higher consumption growth.

The interpretation of these results should be seen with some caution. First, based on the available data, omitted variable bias is minimized but not eliminated. As some of the findings suggest there may be some incidences in which being located in a high risk zone can also indicate other factors that explain living standards but cannot be controlled for with the available data. Furthermore, it is to be highlighted that some of the differences between the results from the *Pooled* and the *Panel* could be due to the different subsamples of households included in these analyses.

2.8. Conclusions

Despite several limitations related to the available data sets, this study reveals important insights into the multi-faceted relationship between poverty and environmental risks based on the combination of high-resolution, geo-spatial data and household data from Vietnam. By using consistent data and methods within the same country, we can shed some light on how the relationship between poverty and multiple dimensions of environmental risk varies as a function of the channels through which poverty and risk interact. The findings highlight that indeed the relationship is hard to generalize, as it depends on the specific types of risks, scale of analysis and locations of interest and could possibly be driven by different linkages.

The main results regarding the four research questions are: (i) at district-level the incidence of poverty is higher in high risk areas, (ii) similarly, at household-level poorer households face higher environmental risks, (iii) for some risks their relationship with household-level consumption varies between rural and urban areas, and (iv) environmental risks explain consumption differences between households, but less so consumption changes over time. In particular, lower household consumption levels are related to steeper slopes, higher rainfall variability and drought hazards in rural communes and higher flood hazards in urban communes. Consumption growth is lower for households in communes with higher air pollution. And for poor households higher rainfall variability and drought hazards relate to lower consumption growth implying the existence of poverty traps as relates to weather risks. Although the data do not allow to conclude that environmental risks generally result in lower consumption growth, these findings suggest that poor people are disproportionally exposed to environmental risks.

More work is needed to confirm the potentially causal relationships between environmental risks and poverty. First, longer time-series data is needed to evaluate whether environmental conditions have long-lasting impacts on household living standards. Second, this study mostly relies on global datasets from remote sensing or modeling work, which needs to be verified and refined with monitored data at the subnational level – optimally from ground stations. Moreover, this study could only define historic risk profiles and not the actual environmental conditions and risk materialization within the survey years. The expansion of these analyses with such data will be an important area for deepening the findings of this work. Furthermore, alternative approaches could be applied to estimate causal impacts: for instance, to estimate deforestation-poverty linkages, future work may use discontinuities provided by borders (either at the sub-national or national level like Crespo Cuaresma *et al.,* (2017) or use instrumental variables such as temperature inversions (which have been used as an instrument for pollution levels as in (Sager, 2019)).

Even in the absence of such analyses, some evidence already exists that shows the detrimental impacts of environmental risks on poverty and household welfare. An accompanying study, for example, shows that actual variation in rainfall and temperature conditions in rural areas is related to significant consumption and income effects also for poor people (Narloch, 2016). In addition, other work has shown the negative welfare and poverty impacts of flood and drought events (Thomas *et al.*, 2010; Bui *et al.*, 2014; Arouri, Nguyen and Youssef, 2015). Moreover, poor people could be affected by environmental risks in many ways that go beyond income and consumption effects, such as detrimental health impacts or a decline in the quality of life due to poor environmental conditions. Such impacts cannot be measured with the available household data.

Altogether, the findings provide important insights of relevance for poverty reduction policies under climate change. The disproportionate exposure of poor people to multiple environmental risks suggests that these people are more vulnerable to climate shocks. At the same time, climate change is likely to increase some of these risks exacerbating already existing poverty traps. Strategies to reduce poverty – especially in rural areas – need to address environmental risks and climate change simultaneously. For example, ecosystem-based adaptation could strengthen ecosystem resilience to climate change and reduce environmental risks, while improving the livelihoods of people depending on these ecosystems. Carefully designed land-use planning policies paired with investments in livelihood support and improved mobility that encourage resettlements and avoid new settlements in high risk areas are another strategy to reduce exposure to environmental risks and climate shocks, while reducing human pressure on already fragile ecosystems. Generally, addressing environmental risks deserves greater attention in poverty reduction policies – especially in the face of climate change.

Chapter 3: Exposure to floods, climate change, and poverty in Vietnam (co-authored with Andrew Smith and Ted Veldkamp)

A version of this article has been published as:
"Bangalore, M., Smith, A. and Veldkamp, T., 2019. Exposure to floods, climate change, and poverty in Vietnam. *Economics of Disasters and Climate Change*, 3, pp.79-99."

Abstract: With 70 percent of its population living in coastal areas and low-lying deltas, Vietnam is highly exposed to riverine and coastal flooding. This paper conducts a "stress-test" and examines the exposure of the population and poor people in particular to current and future flooding in Vietnam and specifically in Ho Chi Minh City. We develop new high-resolution flood hazard maps at 90m horizontal resolution, and combine this with spatially-explicit socioeconomic data on poverty at the country and city level, two datasets often kept separate. The national-level analysis finds that a third of today's population is already exposed to a flood, which occurs once every 25 years, assuming no protection. For the same return period flood under current socioeconomic conditions, climate change may increase the number exposed to 38 to 46 percent of the population (an increase of 13-27 percent above current exposure), depending on the severity of sea level rise. While poor districts are not found to be more exposed to floods at the national level, the city-level analysis of Ho Chi Minh City provides evidence that slum areas are more exposed than other urban areas. The results of this paper provide an estimate of the potential exposure under climate change, including for poor people, and can provide input on where to locate future investments in flood risk management.

3.1. Introduction

Vietnam is a rapidly developing country highly exposed to natural hazards. One of the major natural hazards the country faces is riverine and coastal flooding, due to its topography and socioeconomic concentration: Vietnam's coastline is 3,200 kilometers long and 70 percent of its population lives in coastal areas and low-lying deltas (GFDRR, 2015). Furthermore, climate change is expected to increase sea level and the frequency and intensity of floods, globally and in Southeast Asia (IPCC, 2014; World Bank, 2014). Given the country's concentration of population and economic assets in exposed areas, Vietnam has been ranked among the five countries most affected by climate change: a 1 meter rise in sea level would partially inundate 11 percent of the population and 7 percent of agricultural land (World Bank and GFDRR, 2011; GFDRR, 2015).

Even though climate change impacts are expected to primarily occur in the future, flooding already causes major problems in Vietnam, with some segments of the population more vulnerable than others (Adger, 1999; World Bank, 2010; World Bank and Australian AID, 2014). In particular, evidence suggests poor people are more vulnerable than the rest of the population to natural disasters such as floods, as their incomes are more dependent on weather, their housing and assets are less protected, and they are more prone to health impacts (Hallegatte, Bangalore, *et al.*, 2016, Chapter 3). Poor people also have a lower capacity to cope with and adapt to shocks due to lower access to savings, borrowing, or social protection; and climate change is likely to worsen these trends (Hallegatte, Bangalore, *et al.*, 2016, Chapter 5).

Therefore, it is important to quantify how many people are exposed to floods, how this distribution of exposure falls upon regions and socioeconomic groups, and how climate change may influence these trends. In the spirit of a "stress-test" to examine the seriousness of the issue, this paper employs flood hazard maps and spatial socioeconomic data to examine the following questions in context of Vietnam:

1. How many people are exposed currently? How might this change under climate change?
2. Where is exposure highest currently? How might this change under climate change?
3. How many poor people are exposed currently? How might this change under climate change?

Furthermore, given that the dynamics of poverty and natural disasters (and particularly, floods) occur at the local level, analyses at the national scale (or even at the province or district level) may miss important mechanisms and small-scale differences, from one city block to the next. To complement the analysis across the entire country, we also focus at the local level within Ho Chi Minh City (HCMC), a city with high flood exposure. Here, we combine high-resolution flood hazard data with spatial data on slum location to examine the distribution of exposure across poor and non-poor locations.

While prior studies have examined flood risk in Vietnam, this paper provides two main contributions. First, we develop new high-resolution flood hazard datasets, which incorporate both riverine and coastal flooding and consider climate change[32]. When examining flood exposure, it is important to get as local as possible as impacts can vary widely across space – for instance, impacts can be different from one city block to the next (Patankar, 2015). This paper contributes to the growing literature in economics on the assessment of local and disaggregated disaster impacts (e.g. Del Valle *et al.* (2018) using wind speed), and is the first to our knowledge examining flood risk at this scale in a developing country.

Second, we analyze how flood exposure differs based on socioeconomic dimensions, in this case poverty, across the country and within HCMC. In both cases, we examine how flood exposure and poverty differs spatially across the country (examining the relationship at the district level) but also within HCMC (by identifying slums). The combination of these two datasets on hazard and poverty – typically kept separate in the literature – is another main contribution of this paper.

The consideration of socioeconomic characteristics and the focus on the poor is an important one as not all segments of the population are equally exposed and vulnerable to floods. In some cases, people living in risky places may be richer than the average population: for instance, urban residents are on average wealthier than those living in rural areas (World Bank, 2009). At a more local scale and especially within cities, land and housing markets often push poorer people to settle in riskier areas: where markets factor in hazard risks, housing is cheaper where risk is higher, attracting poorer segments of the population (Lall and Deichmann, 2012; Husby and Hofkes, 2015).

[32] Prior use hazard data at a 1km resolution and are restricted to a single district or city within Vietnam (Apel *et al.*, 2015; Chinh *et al.*, 2017). We develop new high-resolution flood hazard datasets on a 90m x 90m grid.

In addition to differentiated exposure, poor people have higher vulnerabilities: for the same level of losses (e.g. $100), impacts on poor people are much more pronounced than those on richer segments of the population. This poverty-vulnerability relationship operates through the asset, income, and consumption channels. In terms of assets, oftentimes the portfolio of poor people's assets is concentrated in livestock and housing which are vulnerable to floods (Barrett *et al.*, 2013). In comparison, non-poor people tend to have larger amounts of their assets in financial forms (e.g. savings in a bank). Regarding income losses, poor people are more likely to have less diversified income sources (e.g. with income streams reliant on vulnerable assets such as livestock). In comparison, non-poor people's income sources are often more diversified (e.g. including from pensions, capital gains, or remittances). As these income sources are often unaffected by a local flood event, income losses for non-poor people may be less pronounced (Hallegatte, Vogt-Schilb, *et al.*, 2016).

Poor people may also take longer to recover from a flood event due to a limited ability to smooth the shock with limited access to insurance, fewer savings, and limited borrowing capacity. While government support can help after a flood, such support may be inadequate particularly in developing countries with limited technical and financial capacity. In Mumbai, while government support existed after the 2005 floods, it was slow to arrive and the amount delivered was inadequate to support losses suffered (Patankar, 2015). Given that poor people have consumption closer to subsistence, a substantial loss from floods in the absence of support can have high non-monetary costs in the form of irreversible impacts on children and distress sales of assets (De Janvry *et al.*, 2006; World Bank and Australian AID, 2014). In summary, livelihood shocks triggered by floods could keep people from escaping poverty and even push them into deeper poverty (Karim and Noy, 2016).

Despite this poverty-vulnerability relationship, previous quantitative studies on flood exposure in Vietnam have focused on the generation of losses rather than exposure across socioeconomic levels (Ministry of Natural Resources and Environment, 2009; Apel *et al.*, 2015; Chinh *et al.*, 2017). Nevertheless, insights from qualitative focus groups in An Giang, Kien Giang, Kon Tum, Hoa Binh, and Bac Nih find that many poor households feel more vulnerable to floods due to their increased exposure (a result of living in flood prone areas, like along river banks or outside of protective dikes, and often having substandard quality of housing) are less likely to have sufficient assets to buffer the effects of floods, and receive inadequate support to cope (World Bank, 2016a).

This paper provides an in-depth case study of floods, poverty, and climate change in Vietnam and Ho Chi Minh City, examining the exposure of the total population, and poor people in particular to current and future flood hazards, to better understand the problem. While we do not conduct a full probabilistic risk assessment simulating all current and future parameters (which would be extremely challenging), we conduct a stress test to examine how future exposure to floods might look like if current population and poverty trends stay the same. The results from this paper provide new insights on the flood-poverty-climate relationship in Vietnam by examining spatial flood exposure under current and future scenarios. Additionally, these findings provide actionable guidance to help policymakers include socio-economic characteristics (in this case, poverty) when making spatial planning decisions when it comes to flood risk management.

We find at the national-level that a third (33 percent) of today's population is already exposed to a 25 year event, assuming no protection. For the same return period flood under current socioeconomic conditions, climate change may increase the number exposed to 38 to 46 percent of the population (an increase of 13-27 percent above current exposure), depending on the severity of sea level rise. While poor districts are not found to be more exposed to floods at the national level, the city-level analysis of HCMC provides evidence that 68-85 percent of slum areas are exposed to floods, a higher percentage than the rest of the city. The results of this paper provide an assessment of current and future exposure levels, and can provide input on where to locate future investments in flood risk management.

Given these potentially large impacts, a key question of importance to public policy is what is the level of adaptive capacity for households, and whether there are any trade-offs between public response and private preparation particularly in a developing country setting (Goeschl and Managi, 2019). For rural households, recently the Vietnamese Government has increased efforts to improve adaptive capacity and farmers have started to shift cropping patterns (Trinh *et al.*, 2018). However, farmers still have limited understanding of the importance of climate adaptation for their livelihoods and there is a need for better government aid after flood events (Le Dang *et al.*, 2014; McElwee *et al.*, 2017). In urban areas, private adaptive capacity may be even lower, resulting in an increasing need for urban planning to incorporate climate risk dimensions (Liao, Le and Van Nguyen, 2016). While adaptation is not a focus of this paper, the uncertain and growing impacts of flood risk found in this paper indicate more attention is warranted to design effective adaptation policies.

3.2. Data

To examine population and poverty-specific exposure to floods, we employ spatial data defining flood hazard and a number of socioeconomic characteristics representing poverty and population density.

3.3. Flood hazard data

3.3.1. Flood hazard maps for Vietnam developed for this study

For this study, we developed flood hazard maps representing riverine, flash-flood and coastal flood hazards for Vietnam. These flood hazard maps estimate the inundation depth at a grid cell level of 3 arc-seconds, (~ 90m) and provide coastal surge hazard layers, along with pluvial and fluvial layers. The maps provide information on the extent and depth of flood hazard for a specific location. For the coastal component, we explicitly model four return periods – 25, 50, 100, and 200 year events, under current and future climate conditions.

There is a significant amount of uncertainty with regards to how much sea level will rise. For that reason we model three future climate scenarios per return period: a low, medium, and high scenario (Table 15), using estimates from the IPCC (IPCC, 2007, 2014). For the fluvial and pluvial hazards, future climate scenarios were not explicitly simulated owing to the complexity and considerable uncertainties that arise (Smith *et al.*, 2014).[33]

Although robust modeling of the magnitude of future extreme rainfall is not yet possible, heavy rainfall is expected to increase in a warmer climate, owing to the increased water holding capacity of the atmosphere. Therefore instead of a direct modeling approach, future climate scenarios were inferred by taking flood hazard maps derived under current climate conditions for different return

[33] These uncertainties largely arise from climate models; global climate models (GCMs) struggle to represent the physical processes that produce extreme rainfall. Indeed even in higher resolution regional climate models (RCMs), heavy rainfall events are poorly represented. As a result the modelled rainfall data must be 'corrected', in order to render it realistic. The fact that the underlying models themselves cannot represent flood driving rainfall means that there is little confidence in the projections that they produce. Moreover, at the national scale there is very little river gauge data available in Vietnam. Therefore rainfall-runoff models, required to transform rainfall projections into river discharge values, would be largely un-calibrated. This adds an additional source of significant modeling uncertainty to the model cascade. The combination of poorly represented extreme rainfall in climate models, coupled with uncalibrated rainfall-runoff models, would largely render any projections of future flood hazard impractical, owing to the significant uncertainties that arise.

periods, and using them as a proxy for future climate scenarios. The return period hazard maps used for each of the future scenarios are outlined in Table 16. Although simplistic, this method allows areas that may be impacted by increasing riverine and extreme rainfall driven flooding to be identified. Clearly there are some significant assumptions and uncertainties arising from this method. However, given the impracticalities of modeling future flood hazard in Vietnam, this approach provides a plausible and practical attempt to estimate changing flood hazard at the national scale.

For each of the four return periods, four scenarios are modeled (historical, future with low sea level rise, future with medium sea level rise, and future with high sea level rise), combining the coastal and fluvial/pluvial hazard layers (Table 16). For full details on the methodology used to produce these hazard maps, see Appendix C1. Importantly, the flood hazard models do not include flood protection (such as dikes and drainage systems), which can make a large difference in the flood hazard particularly in well-protected areas. In these well-protected areas, our flood maps may overestimate the flood hazard. This is a data limitation that affects many studies of flood exposure: even in high-income countries like the US, flood protection databases are incomplete and many areas are left undefended in national models when in reality they are defended (Wing *et al.*, 2017). While recent work has tried to patch together a database of flood protection (Scussolini *et al.*, 2015), the authors admit much remains unclear and this is a current state of the research.

Table 15. Future scenarios used for Vietnam coastal modeling.

Simulations	Scenario	Percentile	SLR -2100 (m)
Low	RCP 2.6	0.5	0.28
Medium	A1B	0.05	0.6
High	RCP 8.5	0.95	0.98

Note: RCP stands for Representative Concentration Pathway. We use two RCPs from the recent Intergovernmental Panel on Climate Change (IPCC) report (IPCC, 2014) to represent a low climate change and a high climate change scenario. RCP2.6 is a low scenario consistent with temperature increases of 2°C, while RCP8.5 is a high scenario consistent with temperature increases of 4°C. The A1B scenario was taken from a previous IPCC report (IPCC, 2007) and represents a medium climate change scenario, in between RCP2.6 and RCP8.5.

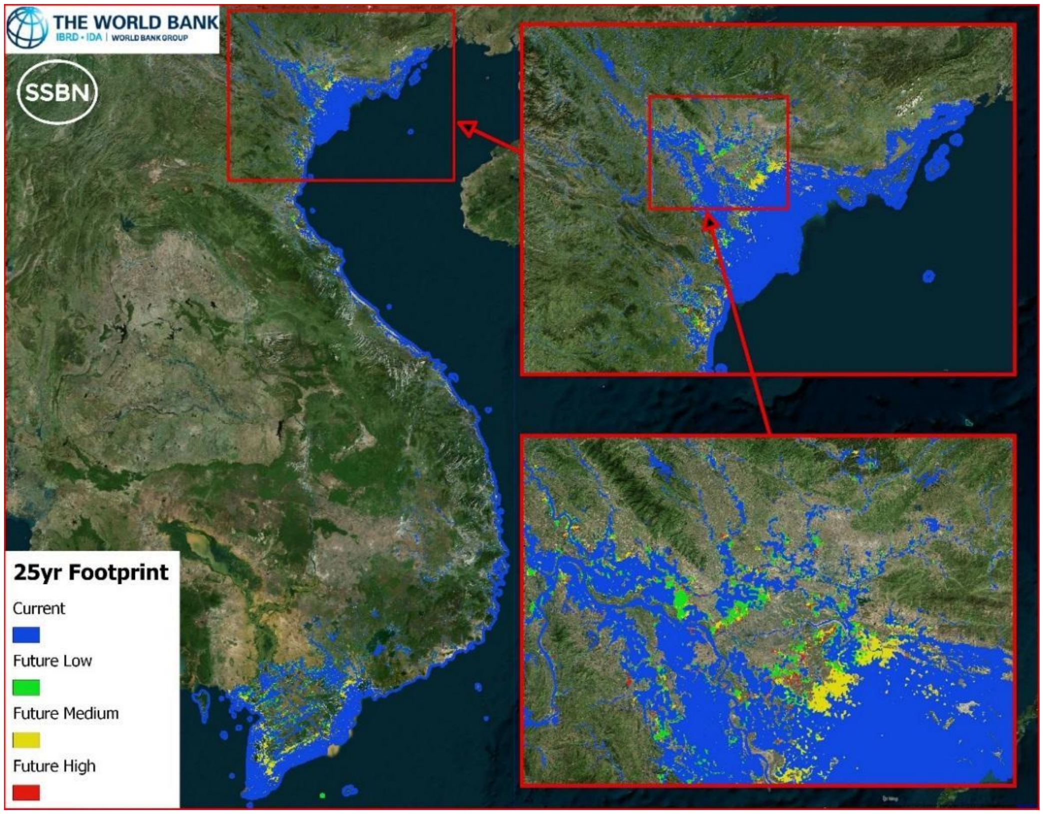
Table 16. Hazard map scenarios for which the modeling was conducted for Vietnam

Scenario	Coastal	Fluvial/Pluvial
1 in 25	1 in 25	1 in 25
1 in 25 Future – Low	1 in 25 + 28cm	1 in 50
1 in 25 Future – Medium	1 in 25 + 60cm	1 in 75
1 in 25 Future – High	1 in 25 + 98cm	1 in 100

1 in 50	1 in 50	1 in 50
1 in 50 Future – Low	1 in 50 + 28cm	1 in 75
1 in 50 Future – Medium	1 in 50 + 60cm	1 in 100
1 in 50 Future – High	1 in 50 + 98cm	1 in 200
1 in 100	1 in 100	1 in 100
1 in 100 Future – Low	1 in 100 + 28cm	1 in 200
1 in 100 Future – Medium	1 in 100 + 60cm	1 in 250
1 in 100 Future – High	1 in 100 + 98cm	1 in 500
1 in 200	1 in 200	1 in 200
1 in 200 Future – Low	1 in 200 + 28cm	1 in 250
1 in 200 Future – Medium	1 in 200 + 60cm	1 in 500
1 in 200 Future – High	1 in 200 + 98cm	1 in 1000

For most of the analyses, the "combined" maps are used, which include both coastal and the fluvial/pluvial floods. For instance, the combined maps for the 25-year return period flood (under current conditions, and low, medium, and high future conditions) are presented in Figure 7.

Figure 7. A visual of what the combined hazard maps (which include coastal and fluvial/pluvial) look like. The map presented here is the worse-case scenario we simulate, a 200-year return period flood with high sea level rise.



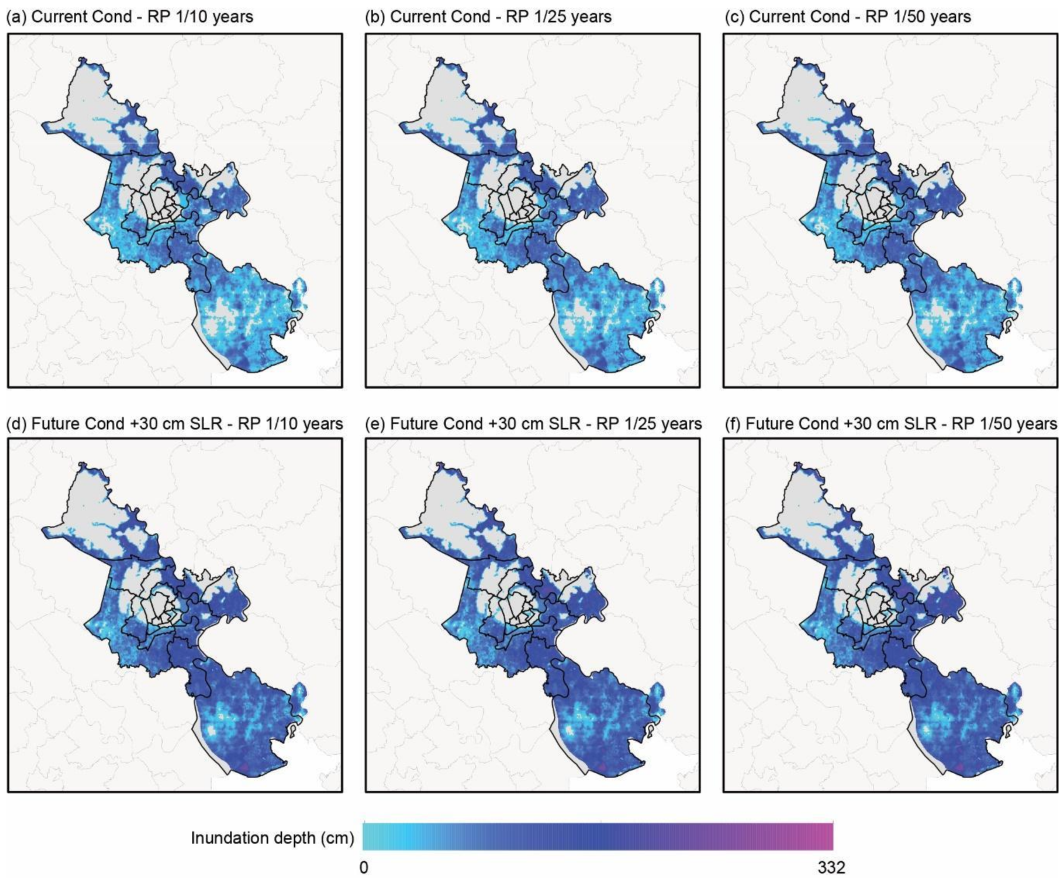
3.3.2. Local flood hazard maps for Ho Chi Minh City

In addition to the flood hazard maps developed for this study as described above, we use an additional set of maps produced specifically for HCMC.

The inundation maps were used in an earlier flood risk study of HCMC (Lasage *et al.*, 2014), and were composed with the MIKE 11 hydraulic modeling software (DHI, 2011). The flood hazard maps, which have a spatial resolution of 20 meters, represent the current conditions for five return periods: 10, 25, 50, 100, and 1000 years. Future conditions, again using the five return periods, include a sea level rise scenario of +30 centimeters in the year 2050 (consistent with the "low" sea level rise used for the maps produced for this study) in combination with current river discharge (FIM, 2013). Potential peaks in precipitation events and/or river discharges due to climate change are not covered by this data set. The inundation layers for a 10, 25, and 50-year return period under current climate conditions and given a sea level rise scenario of +30 centimeters are shown in Figure 8.

Figure 8. Flood maps showing inundation depth (cm) in case of a: (a) 10-year return period flood under current conditions, (b) 25-year return period flood under current conditions; (c) 50-year

return period flood under current conditions; (d) 10-year return period flood given a 30 cm sea level rise; (e) 25-year return period flood given a 30 cm sea level rise; and (f) 50-year return period flood given a 30 cm sea level rise.



3.4. Socioeconomic data

3.4.1. District-level poverty and population data

At the national-level analysis, we overlay the flood hazard maps developed for this study with spatial socioeconomic data. For Vietnam, the World Bank has produced estimates of the number of people within each district who live below the poverty line: this "poverty map" is displayed in Figure 9, panel A and the full methodology can be found in (Lanjouw, Marra and Nguyen, 2013)[34].

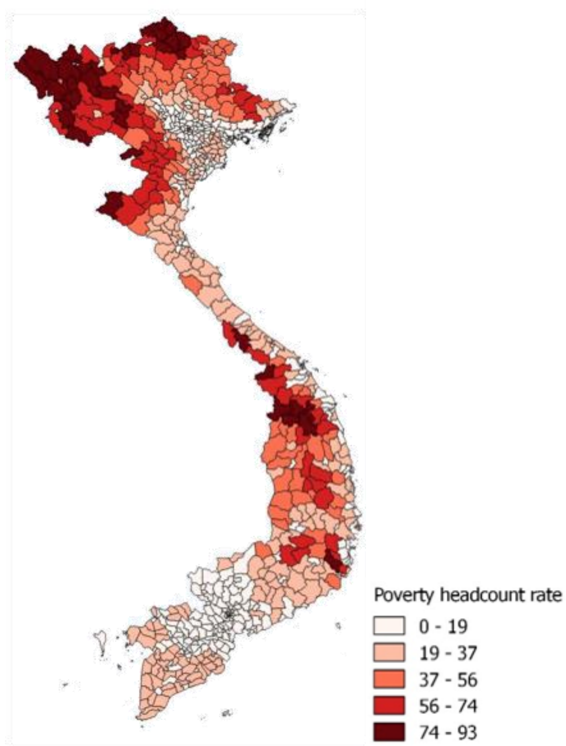
[34] While we considered other metrics of poverty (e.g. the Multi-dimensional Poverty Index), the only available data that was spatially explicit was the headcount and headcount rate, which we use for the analysis. This headcount rate uses an income definition of poverty, with those earning less than $1.25 USD per day classified as poor.

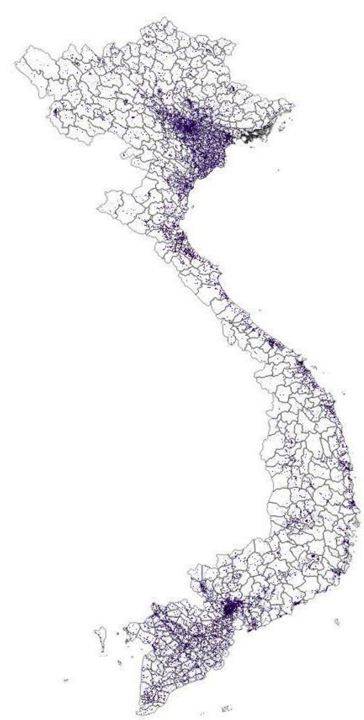
In addition, we use gridded population density data with a 1km resolution from Landscan (Geographic Information Science and Technology, 2015). This "population map" is displayed in Figure 9, panel B.

While we are able to simulate current and future flood hazard, we are unable to project socioeconomic characteristics like poverty or population at the district level, as such an exercise is extremely challenging. Specifically for Vietnam, a recent paper models population and poverty in 2030 at national-level, but notes that examining how these dynamics are distributed spatially within the country is still not possible (Rozenberg and Hallegatte, 2016). Nonetheless, as a stress test, a number of papers in the field have employed the same strategy we follow in this paper, of using current socio-economic characteristics to examine potential future trends, which is considered standard practice (Hirabayashi *et al.*, 2013; Koks *et al.*, 2015; Winsemius *et al.*, 2018).

Figure 9. (a) Poverty map and (b) population density map for Vietnam at the district level.

(A) **(B)**





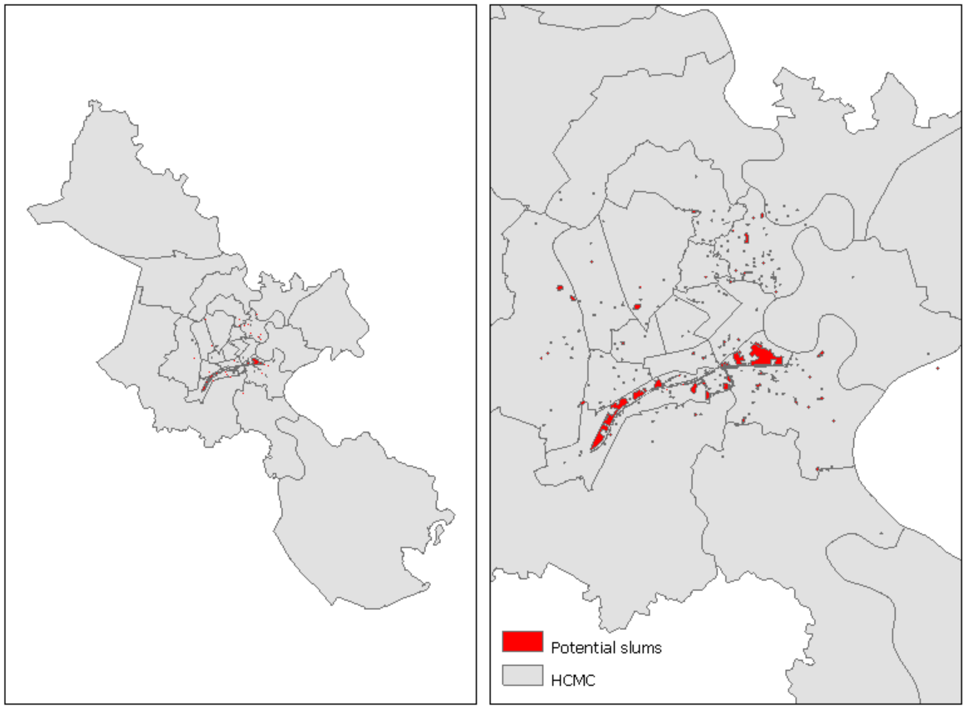
Sources: Lanjouw, Marra and Nguyen (2013); Geographic Information Science and Technology (2015)

3.4.2. Local-level data on urban areas and potential slums in Ho Chi Minh City

The spatial socioeconomic data set used for HCMC is a data set of potential slum areas from 2000 to 2010, from the Platform for Urban Management and Analysis (PUMA), a city-level data set developed by the World Bank (World Bank, 2015b).This data was collected via satellite in the year 2012, through a combination of visual interpretation of various sources and vintages of imagery.

To guide the identification of slums, previous work has provided information on the appearance and geographical extent of slums in HCMC. Surveys of poverty in the city find the appearance of slums in HCMC to be characterized as densely built small households and shelters that have predominantly semi-permanent character (Habitat for Humanity, 2008). In terms of geographic extent, many slums are located in certain districts (districts 2, 3, 4, 6, 8, 11, 12, Binh Thanh, Go Vap, Tan Phu) and along the Saigon River (e.g. Kenh Te, Rach Ben Nghe, Thi Nghe-Nhieu Loc Canal, Kenh Doi, Thi Nghe Canal, Lo Gom, and Canala) (Horsley, 2004; Habitat for Humanity, 2008; De Lay, 2011). Taking into account these spatial and geographic characteristics, the PUMA data set interprets Google Earth imagery to produce two layers of potential slum areas (PUMA, 2013): areas with defined borders (polygon-data) and potential slum areas without (point-data) defined borders. In the latter case, we applied a circular buffer of 50 meters around each point indicating a potential slum location. Evidence suggests that slum areas exist in the northern districts of HCMC (Habitat for Humanity, 2008), which we do not find in PUMA. For this reason, we ran the analyses for two samples – all the districts in the province, and only the districts with potential slums from PUMA. The potential slum locations are presented in Figure 10.

Figure 10. Location of slum areas and locations with urban expansion in the city of HCMC.



Source: World Bank (2015b)

3.5. Methodology

3.5.1. Exposure to flooding at the national level

At the national level, we estimate per district the number of people exposed to each scenario of flooding, and the number of poor people exposed. In the flood data, we define exposed areas as those grid cells where the flood level is greater than zero; non-exposed areas are those grid cells where the flood level is zero. This is a measure of *extent* rather than *depth*, and has been used in previous studies to examine exposure to floods (Ceola, Laio and Montanari, 2014; Jongman *et al.*, 2014; Winsemius *et al.*, 2018). Furthermore, while we lose information by using *extent* rather than *depth* (we have depths in our flood data), we decided to use extent since our flood data assumes no protection. Protection is more likely to impact the depth, rather than the extent, of the flood results.[35]

[35] There is also a good reason for examining extent over depth, in terms of the hazard modeling; flood depths within a large scale flood model are very uncertain, and there is much more certainty about extents.

We then overlay this flood layer with the population density data set, to estimate the number of people per population grid cell that are exposed to floods. As the population density data set is at a lower resolution (1km) than the flood data (90m), we estimate the percentage of the population grid cell which is flooded, and multiply this percentage by the population in that grid cell. For instance, if a population grid cell has 500 people, and 10 percent of that cell is flooded (based on the flood data), then we estimate 50 people to be exposed to floods in that cell. In doing so, we assume that the population is evenly distributed within a grid cell.

We run this analysis for all the scenarios presented in Table 16, and aggregate our results at the district level to estimate the number of people affected. To include the poverty dimension, we use the poverty headcount rate in each district to estimate the percentage of poor people exposed. For instance, if 20,000 people are exposed to floods in District X, and District X has a poverty headcount rate of 20 percent, 1,000 poor people are exposed to floods in that district[36].

3.5.2. Slum exposure in Ho Chi Minh City

For the HCMC analysis, we estimate the general exposure to flooding, for the whole province of HCMC and in each of its 24 districts. The flood maps used here are based on a model of HCMC, and are not the same map as used in the figurative example in Section 3.3.1.

Exposure to flooding was again evaluated using flood extent (we also evaluate flood depth, for full results, see Appendix C2). We examine the flood extent both for all urban areas (the whole HCMC province) and for those areas defined as potential slums (from the PUMA data set) to examine how exposure to floods is different in slum areas.

Again we use a number of events, from the case of regular flooding (10-year event) to more extreme flooding events (1000-year event). Moreover, we examine how this exposure changes due to climate change (proxied by sea level rise changes), by running the analysis with flood hazard maps taking into account a 30 cm sea level rise. In each district and across the whole city, we examine the percentage of area within each of the two categories (all urban areas, and potential slum areas)

---

[36] As evidenced in the HCMC analysis with slum data, poor people are often not evenly distributed but clustered in particular areas. However, due to data limitations across the country, we assume poverty is evenly distributed within a district when conducting the national-level analysis.

that is exposed to floods and the percentage which is not exposed to floods. We then compare these values across the two categories.

3.6. Results

3.6.1. National-level analysis for poverty and exposure to floods

3.6.1.1. Flood exposure (with and without climate change)

At the district level in Vietnam, we estimate the total number of people and the share of the population who are exposed to floods. In the results presented, we examine the four scenarios for the 25-year, 50-year, 100-year and 200-year return period flood – a historical scenario, and three scenarios representing future climate: a low, medium, and high scenario.

We aggregate the results to examine Vietnam's exposure.[37] A third (33 percent) of today's population is already exposed to a 25-year flood in Vietnam, assuming no protection (such as dikes and drainage systems), which can make a large difference in the flood hazard particularly in well-protected areas. In these well-protected areas, our flood maps may over-estimate the flood hazard.

When including climate change, this percentage increases by 13-27 percent, depending on the severity of sea level rise. This increase in exposure is due to the concentration of the population in coastal areas. For the 50-year flood, more than a third (38 percent) of today's population is already exposed. Given climate change, this number is expected to increase by 7-21 percent (resulting in overall exposure of between 40 and 48 percent) for the same return period (50-year). For a 100- and 200-year flood under a high climate scenario, more than half of the population is exposed.

Climate change impacts can be seen in these exposure numbers – for instance, a 50-year flood with medium climate change impacts has the same exposure of a 200-year historical flood (at 44 percent), while almost half the country's population (48 percent) is exposed to a 50-year flood with high climate impacts. Full results are presented in Table 17.

---

[37] Results presented are similar to a previous study analyzing the exposure to a 100-year return period flood without climate change impacts, which finds 40 million people to be exposed to that event (Jongman et al. 2014). While we had planned to compare our simulated results with national statistics, these statistics were unavailable at the time of analysis. However, we consulted our findings with individuals familiar with the Vietnam context who tended to agree that the numbers were in the range of plausible estimates.

Table 17. Population exposed to floods in Vietnam, across the 16 flood hazard scenarios examined.

Scenario	Exposure	Return period			
		25	50	100	200
Historical	Estimated population exposed (million)	30.17	34.30	38.35	40.43
	Percentage of today's population	33%	38%	42%	44%
Low climate change	Estimated population exposed (million)	34.78	36.87	40.91	42.32
	Percentage of today's population	38%	40%	45%	46%
	Increase due to climate change	13%	7%	6%	4%
Medium climate change	Estimated population exposed (million)	38.03	40.22	43.34	45.16
	Percentage of today's population	42%	44%	48%	50%
	Increase due to climate change	21%	15%	11%	10%
High climate change	Estimated population exposed (million)	41.46	43.36	46.13	48.72
	Percentage of today's population	46%	48%	51%	53%
	Increase due to climate change	27%	21%	17%	17%

But not all districts within Vietnam face the same amount of exposure. The spatial analysis also allows us to examine which districts have the highest absolute and the highest relative exposure. We present results for the 25-year flood, for a historical and a high climate scenario (results on geographical extent for other scenarios are similar). For absolute exposure, the largest number of people exposed are found in the Mekong Delta, the Red River Delta, and the Southeast Coast (Figure 11, Figure 12, and Figure 13). But the relative exposure (that is, the percent of the district population which is exposed to floods) shows a larger spread (Figure 14, Figure 15, and Figure 16). Certain regions – including the North Central Coast and the Northeast – have high percentages of their populations residing in flood-prone areas (Figure 14).

Figure 11. Absolute exposure at the district level (total number of people in a district exposed), for a 25-year historical flood (left) and a 25-year historical flood under high climate change (right).

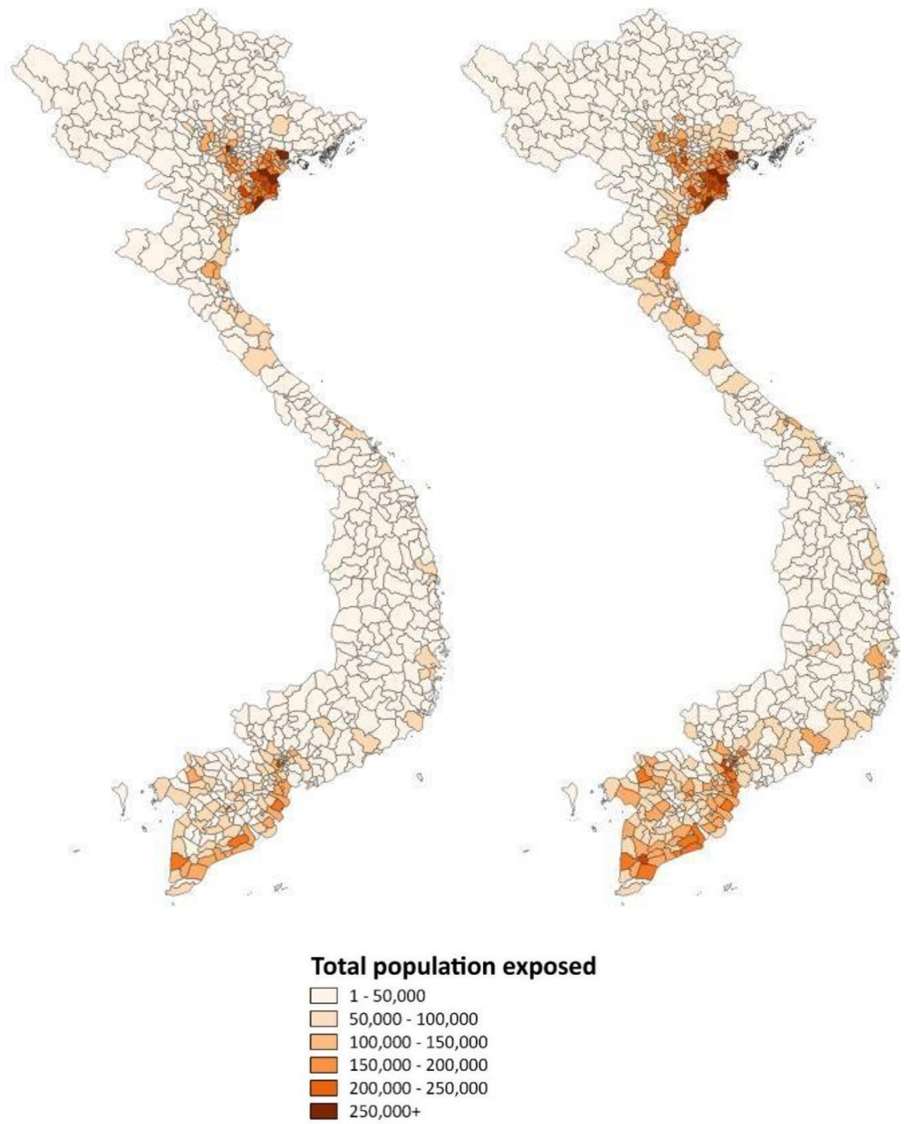


Figure 12. Total population exposed in the Red River Delta for historical 25-year flood (left) and 25-year flood with high climate impacts (right)

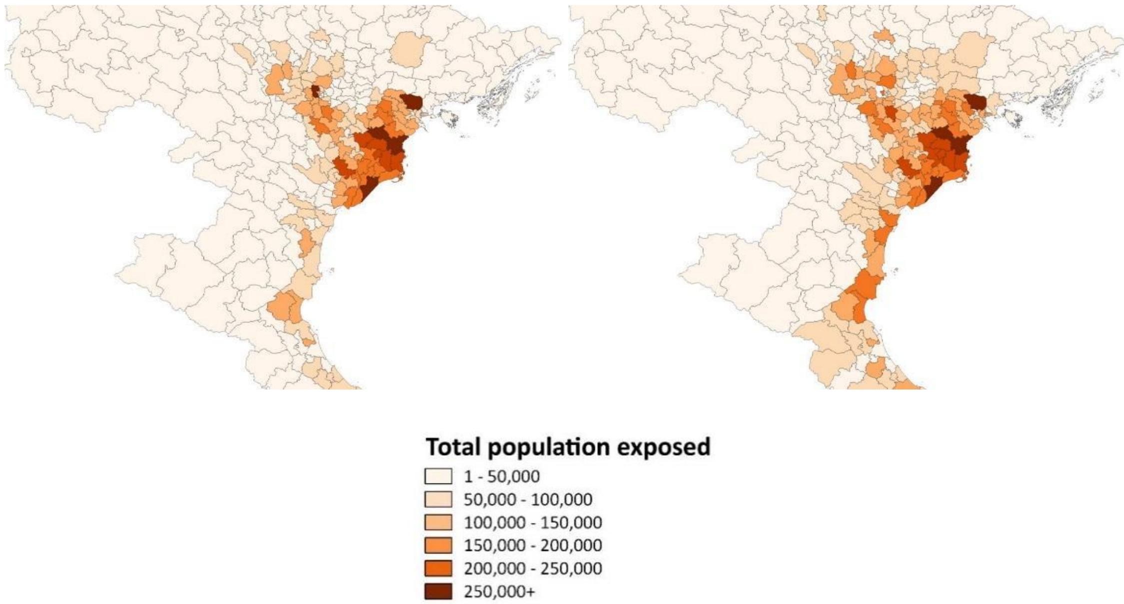


Figure 13. Total population exposed in the Mekong for historical 25-year flood (left) and 25-year flood with high climate impacts (right)

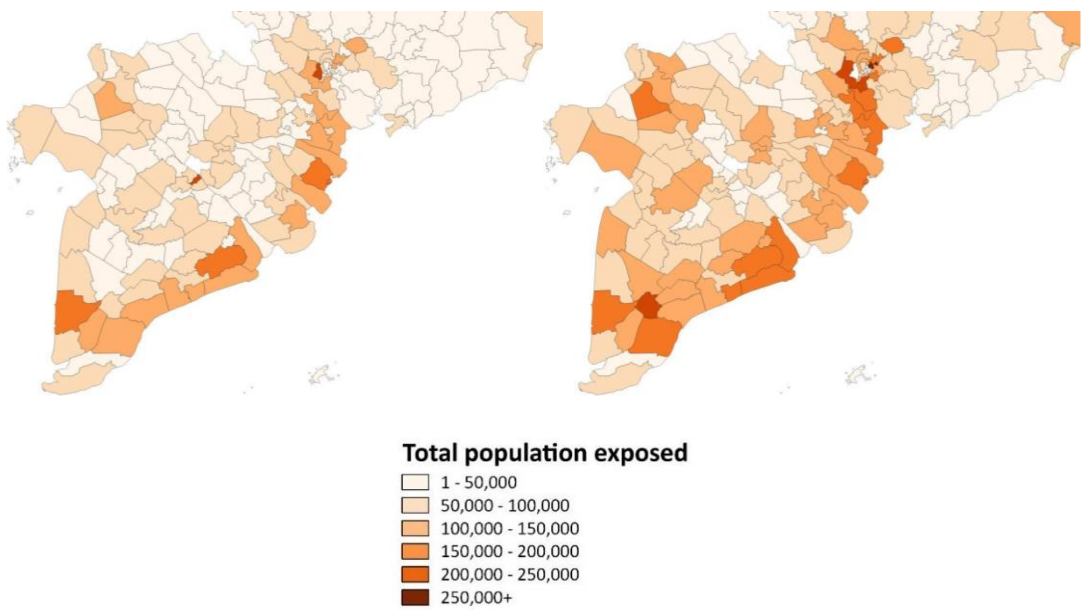


Figure 14. Relative exposure at the district level (percent of district population exposed), for a 25-year historical flood (left) and a 25-year flood under high climate change (right).

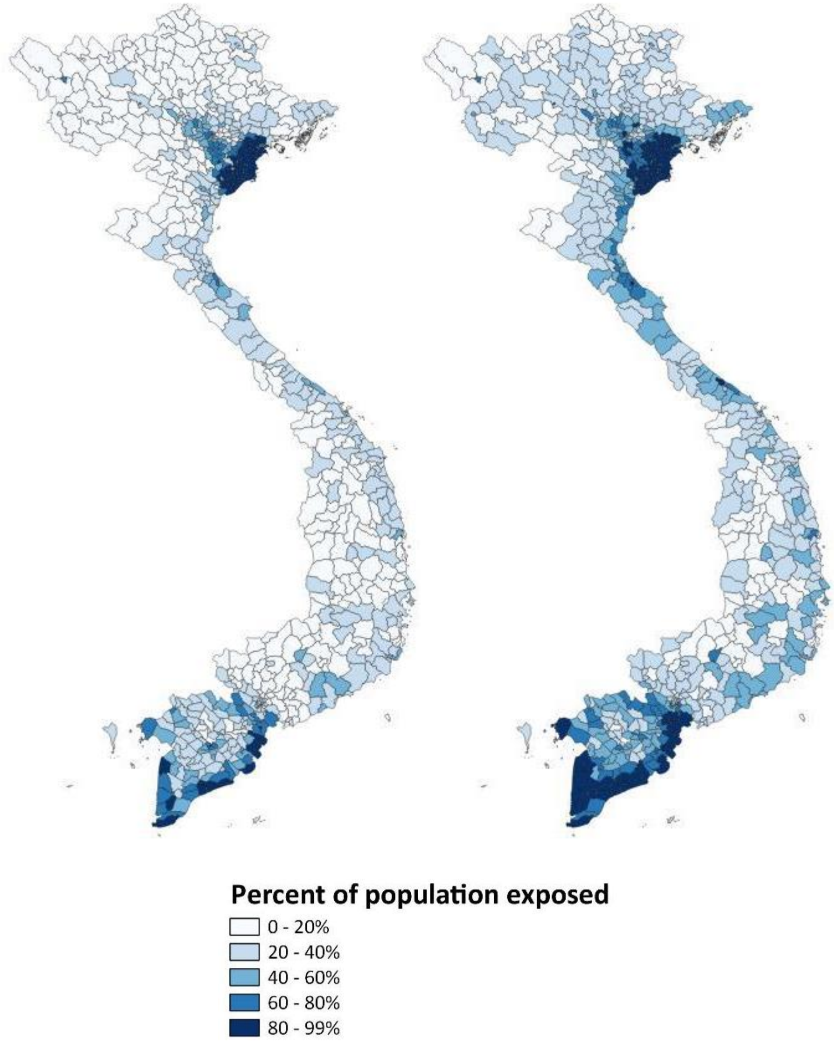
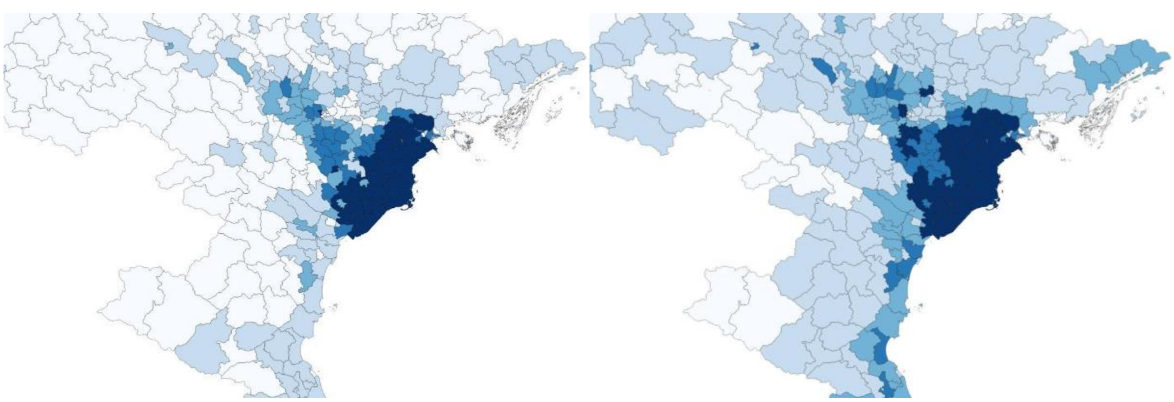


Figure 15. Relative exposure in the Red River Delta for historical 25-year flood (left) and 25-year flood with high climate impacts (right)



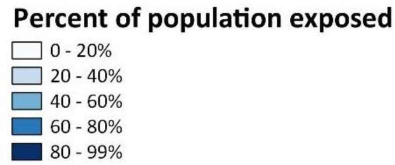
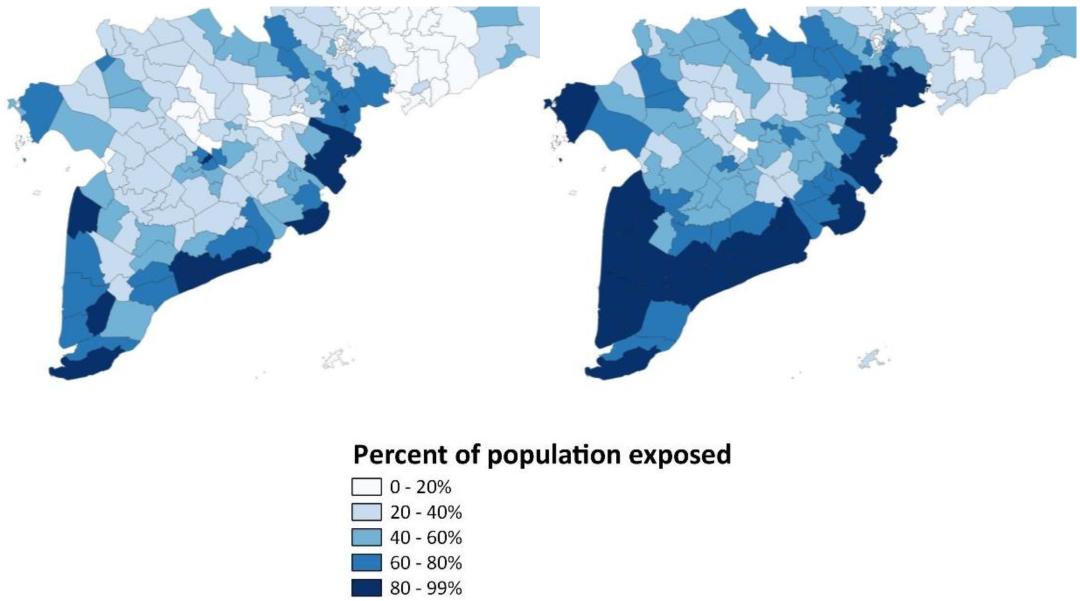


Figure 16. Relative exposure in the Mekong Delta for historical 25-year flood (left) and 25-year flood with high climate impacts (right).



3.6.1.2. Flood exposure and poverty

To examine the question of how many poor people in Vietnam are exposed to flooding, we multiply the population exposure estimates by the district's poverty headcount rate (the percentage of people living below USD 1.25 per day), as calculated in (Lanjouw, Marra and Nguyen, 2013). The results are presented in Table 18.

For a 25-year historical flood, 30 percent of today's poor population is exposed. This number increases by between 16-28 percent given climate change impacts. For a 50-year return period under a high climate scenario, 40 percent of today's poor people in Vietnam are exposed to flooding. For a 200-year return period under a high climate scenario, more than half of today's poor are exposed. Similar to the population analysis, the impact of climate change on the number of poor people exposed is evident. For instance, a 25-year event with high climate change impacts has the same exposure as a 200-year historical event (at around 41 percent of poor people being exposed).

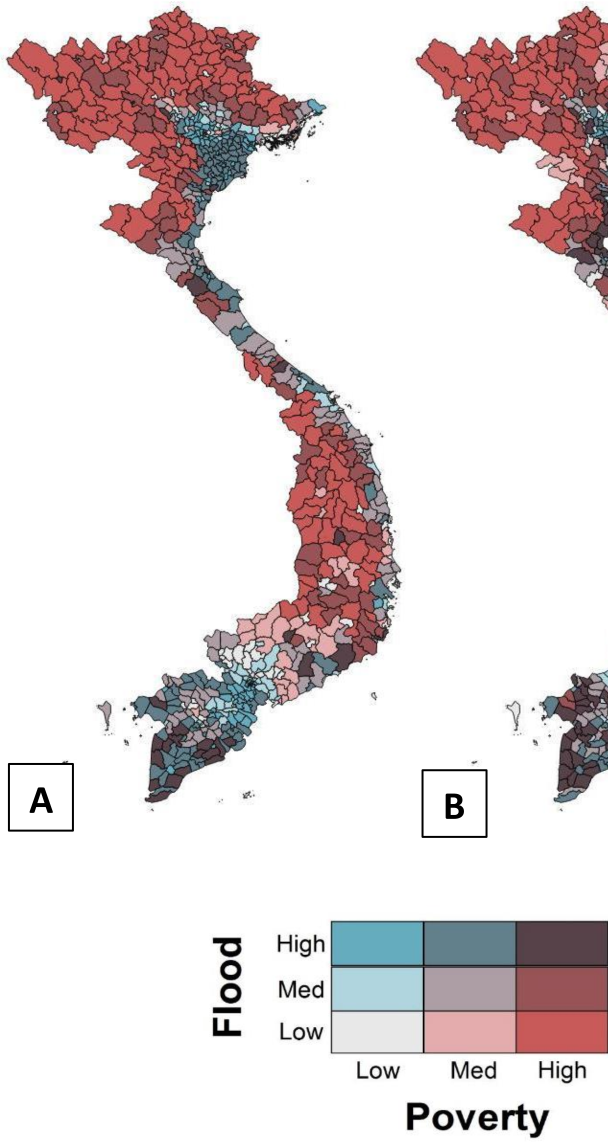
Table 18. Number and percentage of poor exposed to floods in Vietnam, across the 16 flood hazard scenarios examined.

Scenario	Exposure	Return period			
		25	50	100	200
Historical	Estimated poor exposed (million)	5.28	6.19	6.88	7.24
	Percentage of today's poor	30%	35%	39%	41%
Low climate change	Estimated poor exposed (million)	6.27	6.64	7.32	7.54
	Percentage of today's poor	35%	37%	41%	42%
	Increase due to climate change	16%	7%	6%	4%
Medium climate change	Estimated poor exposed (million)	6.80	7.16	7.69	8.00
	Percentage of today's poor	38%	40%	43%	45%
	Increase due to climate change	22%	14%	11%	10%
High climate change	Estimated poor exposed (million)	7.33	7.66	8.14	8.56
	Percentage of today's poor	41%	43%	46%	48%
	Increase due to climate change	28%	19%	16%	15%

Based on the statistics provided in Table 18, there is no strong signal that poor people are more exposed than non-poor people, at the national level. However, this may not be the case in specific regions or within specific districts.

To examine which districts have a confluence of poverty and flood hazard, we classify both each district's poverty headcount rate and flood exposure into three categories: low, medium, and high. We create 3 quantiles for each. We examine both absolute and relative numbers, overlaying the number of poor and number of flood exposed, and the percentage of poor and percentage of flood exposed. The results suggest that areas of the Northern Mountains and the Mekong Delta exhibit districts with high flood and high poverty (darkest shade of brown in Figure 17). The results are slightly different when comparing relative and absolute numbers. When using absolute (the number of poor and number of flood exposed) more areas of high flood and poverty are visible in the Mekong and Red River Delta, as well as along the eastern coasts.

Figure 17. Overlay of poverty and flood at the district level for the 25 year-return period flood with climate change. Map A shows relative exposure, overlaying the percent of poor and percent of population flooded, Map B shows the absolute exposure, overlaying the # of poor and # of population flooded.



Notes: Bins: Map A, Poor, Relative (Low = 0-15%, Med = 15-28%, High = 28%+)
Bins: Map A, Flood Exposure, Relative (Low = 0-26%, Med =26-47% , High =47%+)
Bins: Map B, Poor, Absolute (Low = 0-15,900, Med =15,900 – 31,000, High = 31,000+)
Bins: Map B, Flood Exposure, Absolute (Low = 0-27,000, Med =27,000 – 70,000, High = 70,000+)

However, even though not all of the poorest districts seem to face higher exposure to floods, it is important to remember that poor households and poor individuals within high exposure areas have generally higher vulnerability to the impact of floods. Further, it is very likely that within a district or city, the poorest are the most exposed to floods. We explore this dynamic at the local scale with a city-level analysis of HCMC.

3.6.2. City-level analysis in HCMC for poverty and exposure to floods

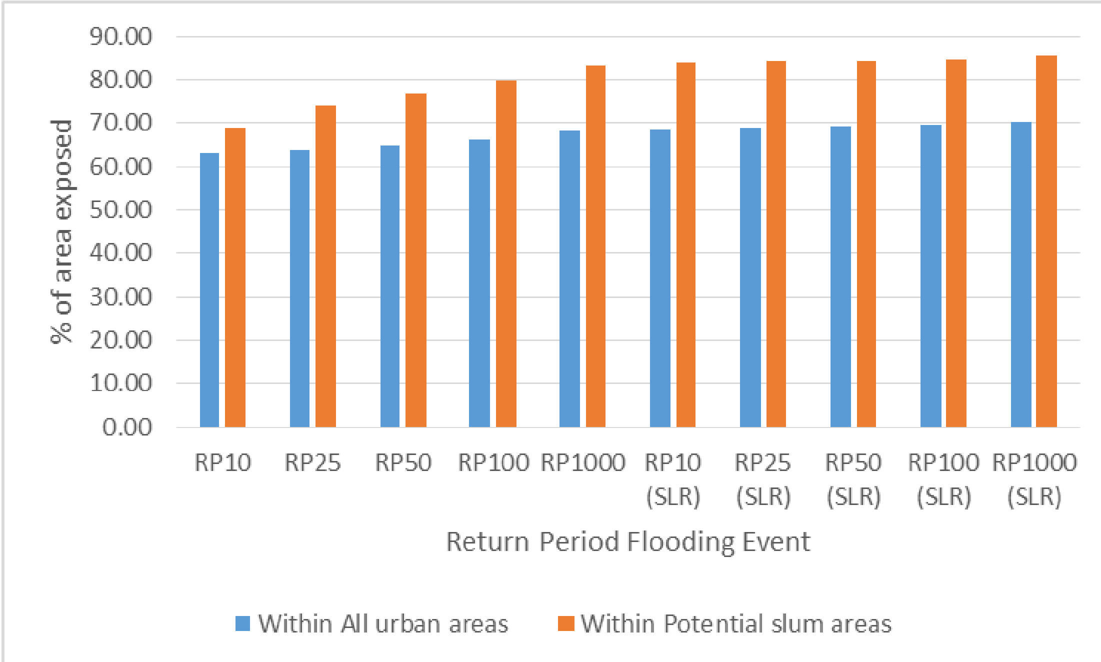
While the relationship between poverty and exposure to floods may not be evident at the national or district level, at a more local scale and especially in urban areas, land and housing markets often push poorer people to settle in riskier areas (Lall and Deichmann, 2012). For instance, comparing exposure of poor people to average exposure, poor households are 71 percent more exposed to flooding in the Mithi River Basin in Mumbai, India (Hallegatte, Bangalore, *et al.*, 2016).

We examine these dynamics in HCMC, using high-resolution local-scale flood maps designed specifically for HCMC (Lasage *et al.*, 2014) and a proxy for poverty using the spatial location of potential slums from the Platform for Urban Management and Analysis (PUMA) data set (World Bank, 2015b). The results we present below are for all districts in HCMC; results for only districts with slum areas are similar and thus not reported.

We find that a relatively high percentage of the potential slum areas are exposed to floods, ranging from 68.9 percent (for a 10-year return period) up to 83.3 percent (for a 1000-year return period). When considering all urban areas of HCMC, exposure to flooding is lower: 63 percent (for a 10-year return period) up to 68.3 percent (for a 1000-year return period). A sea level rise of 30 cm increases the extent of flooded areas the most in slum areas and for a low-probability but recurrent flood (10-year flood). For a 10-year flood and looking only within slum areas, we find an increase in exposure of 15 percentage-points due to sea level rise, compared to a difference of 5.7 percentage-points when looking at the entire urban area of HCMC. These results, as presented in Figure 18, suggest slum areas to be more exposed to floods (and changes in flooding due to climate change) than non-slum areas.[38]

Figure 18. Slum areas tend to be more exposed than the average, across all flood scenarios.

[38] Disaggregated results per district, and results using depth as an indicator can be found in Appendix C2.



Notes: SLR means the scenario includes a 30cm sea level rise due to climate change. RP denotes the flooding events with a particular return period (e.g. RP10 stands for a flooding event with a 10-year return period).

3.7. Discussion and conclusion

This paper conducts a stress-test and presents some initial findings on what exposure to floods looks like in Vietnam, how it may change under a changing climate, and whether poor people are relatively more exposed. Our main contributions are twofold: we first develop a state-of-the-art model at high resolution to represent riverine and coastal flood hazard for Vietnam considering climate change. Second, we examine how exposure to floods differs based on socioeconomic characteristics, with an explicit focus on poverty. The results from this paper provide new insights on the flood-poverty-climate relationship in Vietnam by examining spatial flood exposure under current and future scenarios. Additionally, these findings provide actionable guidance to help policymakers include socio-economic characteristics (in this case, poverty) when making spatial planning decisions when it comes to flood risk management.

Our results indicate that climate change is likely to increase the number of people exposed to floods, especially in the Mekong and Red River Deltas. For the same return period flood under current socioeconomic conditions, climate change may increase the number exposed to 38 to 46 percent of the population (an increase of 13-27 percent above current exposure), depending on the severity of sea level rise. Regarding poverty and exposure, while we do not find evidence of a differential exposure at the national level, we find at the city level that poor people are relatively more exposed to floods. Within HCMC, potential slum areas are 10-20% more exposed to floods compared to the rest of the city, with the exposure differential increasing with climate change.

Nevertheless, the findings presented in this paper should be interpreted considering a number of caveats.

While we use current and future flood data, we only use current population and poverty data, as reliably projecting these socioeconomic trends spatially into the future is almost impossible. Changes in these trends – among many other factors – can lower socioeconomic vulnerability even as the climate change hazard increases (Hallegatte, Bangalore, *et al.*, 2016). Along these lines, while we examine which regions within Vietnam have the highest flood exposure, we do not examine the determinants of vulnerability (other than poverty). Recent analyses suggest that the Northwest, Central Highlands, and Mekong River Delta have the greatest socioeconomic vulnerability (World Bank, 2010).

In the flood hazard maps developed for this paper, we assume no protection due to a lack of data and as a result the hazard maps present an upper bound of flood exposure. Work is currently ongoing to develop a global database of flood protection, and this information can be mobilized for future work (Scussolini *et al.*, 2015). For the national-level analysis, flooded areas are defined as any area with inundation higher than 0. We have not yet explored the depth dimension, although the flood hazard maps developed for this study allow for this potential in future work.

For the HCMC analysis, the location of the slum areas in the PUMA data set are mainly restricted to the old town. Furthermore, slum areas are often difficult to define (with PUMA only identifying potential slums) and the data we have likely does not capture all slum areas within HCMC. In terms of the hazard, the flood maps for HCMC show flood depth and extent from the river and from sea (when looking at the sea level rise scenario). Pluvial flooding and possible 'sink'-areas in the city are not taken into account. Moreover, the lowest return period we have flood maps for is not low, compared to what is experienced in the city. Some areas of HCMC are flooded every year. Since this analysis used a flood with a 10-year return period as the flooding scenario with the highest recurrence interval we were not able to capture the relative differences in exposure to these yearly/bi-annual flooding events (and we hypothesize that poor people are relatively more exposed to these types of flooding than the general population).

Despite these limitations, the analysis presented can offer several points for discussion. The findings of this paper suggest that climate change is likely to substantially increase the number of people exposed to flooding in the future. However, current planning approaches in Vietnam have not yet adequately taken climate change into account (IMHEN and UNDP, 2015). For example, the city of Long Xuyen in the Mekong Delta has based its dike infrastructure around the city on historical floods levels only, with no inclusion of future climate change-induced water levels, despite the poor performance of existing defenses to recent flooding events (World Bank, 2016a). Investments in climate-informed flood protection taken now reduce flood exposure, but can also save money in the long-run by reducing the amount spent on recovery and reconstruction for future floods. And while it is challenging to integrate into project planning, innovative approaches such as decision-making under uncertainty can support policy-makers to design flood projection with climate change in mind (Hallegatte *et al.*, 2012).

This analysis also provides some insights into where to locate flood infrastructure investments when considering socioeconomic characteristics. Generally flood defenses are located in areas where the expected losses are high – which concentrates investments in areas with the highest property and asset values. However, as shown in his paper, socioeconomic considerations such as poverty are important for flood vulnerability. Such a system, prioritizing areas of asset accumulation, will not prioritize poorer places which inevitably have fewer assets. This paper finds that poverty and flood exposure overlap in specific districts of the Mekong Delta and the North Central Coast, which might warrant further attention.

At the city level, we find potential slum areas to be more exposed to floods in HCMC, and that the exposure differential increases with climate change. As a result, risk-sensitive land-use planning may be a priority to ensure development takes place in safer areas. Such planning might encourage development at the outskirts of the city, which are less prone to flood risk, and can be identified from the results of this paper. An important constraint is that for such development to be feasible, they should be paired with transportation investments, which maintain access to the city center where opportunities are presented (World Bank, 2015b).

Despite the potential of risk-sensitive land-use planning, resettlement is the major ex-ante hazard adaptation mechanism employed in Vietnam currently, especially in the Mekong Delta. While such policies can reduce exposure, policy design is critical to ensure the livelihoods of the poor are supported. For instance, surveys in Tan Chau district suggest the resettlement policy enacted in

2002 may have made households worse-off: inadequate financing resulted in households paying for their new settlements out-of-pocket; many households who were farmers and fishers did not have adequate land, transportation and market access, and inadequate livelihood support was provided to them (World Bank, 2016a). Where resettlement policies are enacted, it is imperative that such policies are paired with livelihood and financing support.

Beyond strategies to reduce exposure, other policy options to reduce vulnerability to improve households' ability to adapt may warrant increased attention. Strategies such as government subsidies for household-level flood protection (like raising of floors), improved financial inclusion, and better observation systems and early warning, and resilient agricultural practices can reduce the asset and income losses associated with floods (Hallegatte, Bangalore, *et al.*, 2016). When hit, targeted social protection (which can support the affected population quickly after a large flood) can hasten recovery (Hallegatte, Bangalore, *et al.*, 2016, Chapter 5). Such policy measures may be targeted in areas with higher future exposure (geographical targeting) as well as to individuals and households classified as poor and near poor who experience flooding (individual targeting). Areas such as the Northern Mountains have high poverty and are expected to experience an increase in flood exposure. While infrastructure protection can be costly in these remote and sparsely-populated areas, strategies to reduce vulnerability or improve the ability-to-adapt of households can reduce flood impacts.

The results of this paper provide an estimate of the potential exposure under climate change, including for poor people, and can suggest increased attention and investments be directed towards improving adaptive capacity. Future research on how to design such policies, and how to enable institutional framework conditions to enable private adaptive capacity may be a priority for future research.

Chapter 4: Household exposure, vulnerability, and ability to respond to Nigeria's 2012 floods

Abstract: Nigeria experienced severe flooding during the rainy season of 2012 which cost $17 billion USD in macro-economic terms. This paper examines the local impacts of the flood on livelihoods, which can be severe but are often overlooked in aggregate statistics. Combining panel household survey data before and after the flood with satellite imagery, I provide evidence on the exposure, vulnerability, and ability to respond of households to better understand flood risk at the local level in a sub-Saharan African context. This paper employs multiple modalities of measuring exposure to isolate the causal impact of the flood and uncover local spillovers and interactions with market factors. On average, I find affected farmers lost around 20% of crop production and 50% of value. Additionally, impacts depend on the interaction between the individual and area-level measures of flood exposure. While farmers who live in a flooded area and directly experience the flood lose half their crop value, those in flooded areas but not directly affected experience increases in crop value by around 30%, which is related to food price increases. I also find that households living outside of historically flooded zones experience more severe impacts. In terms of household responses, I do not find evidence of consumption or livestock changes for the full sample, but poor affected households do experience drops in consumption and livestock value. The quantitative results across the country, combined with evidence from the village of Osomori suggests that households face several constraints to their recovery and may not receive adequate support to foster adaptation. The findings of this paper provide a more nuanced and comprehensive understanding of the local impacts of floods in sub-Saharan Africa and can help inform disaster risk management strategies in an era of increasing risk due to climate change.

4.1. Introduction

The continent of Africa accounted for less than 1% of total economic losses from natural disasters during the time period of 2001-2020 (UC Louvain, CRED and USAID, 2022). Despite this low global share, flood impacts in sub-Saharan Africa are sharply rising both in terms of the population exposed to floods (Tellman *et al.*, 2020) and vulnerability measured in fatalities (Di Baldassarre *et al.*, 2010; Tanoue, Hirabayashi and Ikeuchi, 2016). Socio-economic factors also interact with risk: almost 75 million people in sub-Saharan Africa face both high flood risk and extreme poverty (Rentschler, Salhab and Jafino, 2022), more than any global region. Due to this high exposure, vulnerability, and lower capacity to adapt due to resource constraints, floods in the region have the potential to roll back decades of progress on poverty reduction and development (Hallegatte, Vogt-Schilb, *et al.*, 2016; Hallegatte *et al.*, 2020).

Despite the severity of impacts and rising vulnerability in the region, flood impacts in sub-Saharan Africa remain understudied in the literature (Botzen, Deschenes and Sanders, 2019). Additionally, most existing studies in the region explore aggregated, macro-economic impacts of floods while the recent literature finds that most of flood impacts are highly localized and depend on geography and socio-economic vulnerability (Felbermayr *et al.*, 2022). Factors such as population distribution (especially the distribution of marginalized groups), settlement in floodplains, economic structure and reliance on rain-fed agriculture, and policy frameworks on disaster risk management all interact to determine flood impacts locally (Koks *et al.*, 2015; Kocornik-Mina *et al.*, 2020).

One country particularly at risk is Nigeria, which experienced severe floods in 2022, 2016, and 2012 (Al Jazeera, 2022), and this paper explores the impacts of the 2012 event. From July to October of 2012, the country experienced unprecedented floods which affected 30 of its 36 states, led to more than 431 fatalities, displaced around 1.3 million people, and cost $16.9 USD billion in damages (Unah, 2017). While damage estimates describe the large aggregate costs to the economy, they may not capture the full impacts on livelihoods, or inform about their distribution (Hallegatte, Vogt-Schilb, *et al.*, 2016).

This paper examines the exposure, vulnerability, and ability to respond of households to Nigeria's extreme flood event in 2012 and makes three main contributions. First, I provide a more comprehensive accounting of flood impacts in sub-Saharan Africa by exploring local household outcomes across Nigeria for each component of the risk framework: exposure, vulnerability, and ability to respond. Second, methodologically, I combine satellite imagery and household self-reported data to estimate the causal impact of the flood on agricultural and socio-economic outcomes using a difference-in-difference setup. I also combine this quantitative analysis across the country with evidence from an interview with an individual impacted in the village of Osomori in Anambra state was severely impacted in 2012. Third, I examine spillovers and interactions with local market factors which uncover nuanced relationships between households and flood risk in a sub-Saharan African context.

By combining household survey data with spatial data before and after the flood, I find affected farmers lost around 20% of crop production and 50% of value on average. Additionally, impacts depend on the interaction between the individual and area-level measures of flood exposure. While farmers who live in a flooded area and directly experience the flood lose half their crop value, those in flooded areas but not directly affected experience increases in crop value by around 30%, which is related to food price increases. I also find that households living outside of historically flooded zones experience more severe impacts. In terms of household responses, I do not find evidence of consumption or livestock changes for the full sample, but poor affected households do experience drops in consumption and livestock value. This quantitative evidence across the country is complemented with an interview conducted with an individual from the village of Osomori and suggests finds that households face several constraints to their recovery and may not receive adequate support to foster adaptation. Given the increases in precipitation expected in the future linked to climate change (IPCC, 2018), the results from this paper studying the 2012 flood can help inform future decision-making surrounding disaster risk management strategies in Nigeria and across The Global South.

4.2. Literature and Study Context

4.2.1. Empirical evidence on the economic impacts of natural disasters

The literature on the economic impacts of natural disasters is extensive, and has been studied through theoretical models, computable general equilibrium models, and more recently, through empirical analysis (Botzen, Deschenes and Sanders, 2019). Almost all of the more than 100 existing empirical studies explore the impact of natural disasters on economic outcomes (e.g. gross domestic product), using a panel data regression framework (Klomp and Valckx, 2014; Lazzaroni and van Bergeijk, 2014; Cavallo, Becerra and Acevedo, 2022). One consistent finding emerging from these studies is that natural disasters have the largest economic impact on GDP in low and middle-income countries (Klomp and Valckx, 2014; González, 2022) and especially in sub-Saharan Africa (Adjei-Mantey and Adusah-Poku, 2019; Cavallo, Becerra and Acevedo, 2022).

The empirical setup in almost all of these studies uses panel data aggregated at large geographic scales (e.g. country-year, province-year, or county-year level), which often fail to consider that most natural disasters have localized impacts (Botzen, Deschenes and Sanders, 2019). Who a natural disaster hits, where it hits and with what intensity all matter and determine the direct and indirect impacts on the ground. Recent global analyses use data at the disaggregated local level using advances in detecting economic activity from space (e.g. night-time lights), and find that natural disasters do reduce local economic activity especially in developing countries, but impacts are not long-lasting (Kocornik-Mina *et al.*, 2020; Felbermayr *et al.*, 2022). Felbermayr *et al.*, (2022) documents positive spatial spillovers as neighboring grid cells take over economic activity lost in an affected area; while Lima and Barbosa (2019), in an analysis of floods in Brazil, document negative spillovers with neighboring (but unaffected) areas reducing economic output.

The results from this literature show that impacts of weather shocks are a highly localized phenomenon, in terms of economic outcome studied, representation of the hazard, and the underlying mechanisms through which spillovers propagate. In outlining steps for future research, Felbermayr *et al.* (2022) suggest exploring how economic factors, households, and firms react to floods at the local level to better understand impacts and formulate policy responses, especially in low-income countries, where 90% of natural disasters occur (Klomp, 2016).

4.2.2. Exposure, vulnerability, and ability to respond to floods at the local level in low and middle-income countries

Flood impacts on human communities depend on the intensity of the hazard, the exposure of the population, existing social vulnerability, and the ability of affected communities to respond (Cutter, 1996). Households across The Global South, and especially the poorest[39] within those countries, tend to be more exposed, more vulnerable, and less able to adapt with natural disasters pushing

[39] Each study reviewed in this paper has different definitions of who is “poor” and who is “non-poor”. In general, most measure poverty monetarily using expenditure or income data, and classify the poorest as those in the bottom 40 or bottom 20 percent of the distribution.

an estimated 26 million people into poverty every year (Hallegatte, Vogt-Schilb, *et al.*, 2016; Hallegatte *et al.*, 2020). While the literature on drought impacts in sub-Saharan Africa is extensive (e.g. (Kazianga and Udry, 2006)), the literature on the local impacts of floods and storms across The Global South is considerably smaller but growing and is the focus of this section. Many of these studies use micro-level data from panel household surveys before and after an extreme event. Furthermore, as more recent household surveys contain geographical representation, the location of the household can be merged with satellite data on floods and weather shocks. These local analyses provide a more precise understanding of who was exposed, how vulnerable households were, and their ability to respond in the years after the shock.

In low and middle-income country contexts, households may settle in floodplains where land is cheaper, water can be diverted for agricultural production and for transport connections, which increases exposure to riverine flooding (Hallegatte *et al.*, 2020). Additionally, people in low-income countries are significantly less protected from flood infrastructure compared to high-income countries (Scussolini *et al.*, 2015). This difference in protection alone can explain a factor of 100 in flood risks between poor and rich countries before population vulnerability is considered. A recent global study finds that low- and middle-income countries are home to 89% of the world's flood exposed people; and of the 170 million facing high flood risk and extreme poverty, almost half are in sub-Saharan Africa (Rentschler, Salhab and Jafino, 2022). The difference is also true within countries: investments —including those in disaster risk reduction—are often directed toward the relatively wealthier areas at the expense of poorer neighborhoods. Within countries, a global analysis of 52 countries across The Global South finds that poor households in sub-Saharan Africa tend to be more exposed to riverine flooding, especially within urban areas (Winsemius *et al.*, 2018). Case studies of Dar-es-Salaam, Tanzania and Mumbai, India find a similar over-exposure of poor people when examining specific flooding events (Patankar, 2015; Erman *et al.*, 2019).

However, the relationship between poverty and exposure to floods varies based on region and spatial scale, and historical factors. In the countries across Latin America and South Asia studied by (Winsemius *et al.*, 2018), poor households are often less exposed to floods. In Vietnam at the district and commune level, areas with low poverty rates have three times higher flood exposure than those with high poverty rates due to the higher incidence of flood hazards in prosperous coastal areas and river deltas (Narloch and Bangalore, 2018). At the household level, poor households in rural areas in Vietnam are similarly more exposed; however, in urban areas, poor households are more exposed, supporting a finding from a separate study on Ho Chi Minh City (Narloch and Bangalore, 2018; Bangalore, Smith and Veldkamp, 2019). Yet, the urban bias is not confirmed across all developing country cities: in Accra, Ghana, both poor and non-poor households are equally exposed to floods (Erman *et al.*, 2020).

Furthermore, the relationship between floods and poverty also depends on reporting biases when relying on self-reports. Most studies in the literature assess flood exposure by overlaying household location data with hazard maps; however, some papers also ask households to self-report exposure. The question of how long people can reliably remember an item of information is unknown, as very few psychological studies of memory have dealt with timescales longer than a year (Fanta, Šálek and Sklenicka, 2019). Emotional events such as a natural or human-caused disaster are more likely to be remembered (Berntsen and Rubin, 2006), but memories of traumatic events can become inconsistent within one year (Hirst *et al.*, 2015). Depending on how long ago households are asked to recall a flood event, they may or may not provide accurate information. Households may also systematically over-report information in household surveys if they believe responding in a certain way may be tied to assistance from the surveyor, government or NGOs (Dillon and Mensah, 2021).

Self-reported exposure to a specific flood event also depends on a household's historical experience. People may adapt to the average flood exposure conditions that they have experienced, and use them as a reference to interpret deviations from that average (Guiteras, Jina and Mobarak, 2015). Evidence from Bangladesh suggests that households who live in areas with lower historical flood levels are more likely to report experiencing the 2004 event as a "flood" as the deviation between the 2004 flood size and the average flood size grows. However, for high historical exposure households, the likelihood of reporting a "flood" is inelastic to flood size (Guiteras, Jina and Mobarak, 2015). The authors of this study suggest these biases render "self-reports of little value, and points to the need for objective measures". However, "objective" measures of flood experience – usually captured through remote sensing to delineate flood zones, are also subject to measurement error, especially as many satellite products currently used cannot see through cloud cover (NASA, 2022).

While it is not always clear that poor households are always more exposed to floods, households in low- and middle-income countries tend to be more vulnerable when hit, losing a larger share of wealth compared to individuals in high-income countries (Beegle and Christiaensen, 2019). One reason is asset vulnerability: poor people tend to have less diversified portfolios and their savings are more vulnerable to natural hazards (e.g. livestock and crops) compared to higher-income individuals who hold a larger share of their wealth in intangible savings such as bank accounts (Hallegatte, Vogt-Schilb, *et al.*, 2016). Furthermore, the quality of assets owned by poor people tends to be more vulnerable to climatic shocks. Households living in slums or informal settlements constructed of wood, bamboo, and mud and occupying steep slopes tend to suffer more damage from a flood or storm than households whose homes are made of stone or brick (Akter and Mallick, 2013; Patankar, 2015).

In rural areas without functioning markets, poor households are highly dependent on agricultural income and ecosystems and are therefore vulnerable to the impacts of floods on yields and the health and functioning of ecosystems. While "normal" floods during the rainy season can be beneficial for area under cultivation and agricultural productivity (Banerjee, 2010), "extreme" floods are likely to submerge cropland and destroy the harvest. For example, Cyclone Nargis hit Myanmar in 2008 and killed 140,000 people, led to widespread flooding, decreased yields and lowered agricultural incomes (World Bank, 2015a). After the 1998 Great Floods in Bangladesh, agricultural households experienced a 42-62% drop in crop production (Del Ninno, 2001). After Cyclone Aila hit the same country in 2009, extensive crop damage and livestock losses reduced incomes for agricultural households by 33% (Akter and Mallick, 2013). While natural capital can serve as a safety net after a disaster (Barbier, 2010), depleted ecosystems and climate change impacts may impair their ability to smooth consumption in the face of shocks (Wunder, Noack and Angelsen, 2018).

To compound vulnerability, households in developing countries spend a large share of their household budget on food, and agricultural supply reductions due to floods can increase food prices and impact health. Evidence suggests that poor people in developing countries spend between 40-60% of their overall budget on food – far more than the 25% spent by nonpoor people (Hallegatte, Bangalore, *et al.*, 2016). This has consequences during shocks: a study in Mexico by Rodriguez-Oreggia *et al.* (2013) find that floods increase food poverty by 3.5% at the municipal level. In Pakistan in 2010, unprecedented floods destroyed 2.1 million hectares of agricultural land, decimating production and reserves which resulted in a 50% increase in the price of wheat (Cheema *et al.*, 2015). Also in 2010, Tropical Storm Agatha increased poverty by 16% and reduced per capita consumption by 7.7% for households in Guatemala, with impacts most severe for the food consumption of urban households (Baez *et al.*, 2017). The authors provide evidence that a 6-27%

increase in food prices for staple goods in urban areas led to consumption decreases, which underlines the importance of understanding how the impacts of weather shocks on production in rural areas potentially propagate to urban areas.

Not all studies document increases in food prices after large floods. In Mozambique, Baez, Caruso and Niu (2020) find that the 2006 flood left maize market prices largely untouched and there was no impact on the total consumption of households. The authors suggest that safety nets operated efficiently in this context allowed households to maintain food consumption. But even when total food expenditures remain unaffected, calorie consumption and expenditures on calorie-dense foods can fall, as experienced in Bangladesh after the 1998 Great Flood (Del Ninno, 2001). Food security consequences from weather shocks also interplay with market access: in the Philippines, (Safir, Piza and Skoufias, 2013) found a 4% decrease in food consumption in areas with low precipitation, but this effect disappears in areas close to highways.

More generally, poor people are less able to cope with income losses by adjusting their consumption basket. They cannot cut back on luxury consumption or delay consumption the way wealthier households can. In many countries they are close to the subsistence level, which means that reducing consumption can have immediate negative impacts on health (if food intake is reduced or medical care becomes unaffordable), education (if children are taken out of school), or economic prospects (if essential assets have to be sold) (Karim and Noy, 2016). These consumption cuts have a large impact on immediate well-being, but can also affect human capital through health or education, creating long-term consequences on income and prospects.

Anttila-Hughes and Hsiang (2013) provide a case study to illustrate the difficult choices poor households make in response to typhoons in the Philippines. The authors find that both rich and poor households lose on average 6.6% of income and 7.1% of expenditures, and that these local impacts in the aftermath of typhoons can be 15 times larger than the economic damages associated with the event itself. Households were unable to mitigate storm-induced losses via consumption smoothing strategies like transfers, savings, and borrowing, and spent less on medicine, education, and nutrient-dense foods. The authors calculate that these behaviors in the aftermath of large storms had tragic consequences for children, as typhoons are responsible for 13% of overall infant mortality across the country.

Another aspect to consider when exploring vulnerability to floods is the spatial scale of existing empirical work and differential impacts in rural and urban areas. For example in Vietnam, Arouri, Nguyen and Youssef (2015) finds that floods reduce consumption by 4.4% at the commune level. Building on this work, Narloch and Bangalore (2018) identify that the relationship between flood hazards and consumption differ for rural and urban communes in Vietnam. A significant negative impact on consumption is only detected in urban areas, while the relationship between floods and consumption is close to zero for rural areas. In this context, crops such as rice can withstand heavy rainfall and support rural incomes, while flooding in cities disrupts infrastructure and people's ability to get to work and earn an income (World Bank and Australian AID, 2014). Even at the city scale across sub-Saharan Africa, inconsistent results can be found from two studies of Dar-es-Salaam, Tanzania and Accra, Ghana (Erman *et al.*, 2019, 2020). In Accra, poorer households lost a much larger share of their income to floods, while no relation between losses and poverty is found in Dar-es-Salaam.

The focus of this paper is on the local impacts of floods in sub-Saharan Africa, which remains under-studied in the literature despite the increasing exposure and vulnerability for households in the region (Botzen, Deschenes and Sanders, 2019; Hallegatte *et al.*, 2020; McCarthy *et al.*, 2021). The two studies that are most closely related to this paper are in Malawi (McCarthy *et al.*, 2018) and in Tanzania (Tiague, 2021). McCarthy *et al.* (2018) analyze the impacts of floods that occurred during the 2014/15 growing season in Malawi and find that maize yields and the value of crop production decreased by half. When looking at impacts on consumption, on the surface it appears as if households were able to withstand the shock: drops in food consumption expenditures were less dramatic, and calories per capita were higher. However, dietary diversity was 25% lower in highly-affected areas, suggesting that households adjusted their consumption basket to maintain total intake at a lower cost at the expense of micronutrients. Tiague (2021) document a 34% drop in crop production due to floods in Tanzania and find persistent effects even 2 years later, and similarly find no impact on total household consumption.

This paper builds on these influential studies to contribute to the literature on flood impacts in sub-Saharan Africa. While both papers explore impacts on agricultural households, I examine how agricultural impacts propagate to both rural and urban households. Additionally, I explore flood impacts on a range of crops grown instead of focusing on just one crop (McCarthy *et al.* (2018)) or using aggregated production data (Tiague (2021)).

In terms of defining floods, I use spatial data that may more accurately represent flooding, combine this information with self-reported data, and use existing hazard maps to examine historical exposure. For flood definition, McCarthy *et al.* (2018) use a combination of rainfall and household specific variables like elevation, which may not adequately represent flood impacts at the local level (Guiteras, Jina and Mobarak, 2015). Tiague (2021) uses data from the Dartmouth Flood Observatory; while commonly used in the literature, the areas that are flooded are represented using large polygons (tens or hundreds of square kilometers) and may not be accurate for defining household exposure. Furthermore, I combine spatial data with self-reported data in the same context to provide an understanding of what each measure represents and how these variables interact. I also use data from flood hazard maps to understand which households lived in historically flooded zones, and compare outcomes for households who lived outside of hazard zones but were affected by the event of study in 2012.

In addition to exposure and vulnerability, poor people often have fewer resources and struggle more to cope with and recover from disasters. Poor people tend to be more alone: they receive little support from financial instruments, social protection schemes, and tend to have less power in decision-making. In developing countries, poor people are usually unable to access disaster insurance due to affordability issues, large transaction costs for small portfolios, weak institutions, and a lack of trust (Kunreuther, Pauly and McMorrow, 2013). Poor people in developing countries are often excluded from governance, which reduces their ability to access post-disaster support. Aldrich (2012) argues the poorest are disadvantaged before and after a disaster across multiple dimensions: some programs are tied to formal employment, while most poor people work in the informal economy; poor people in remote rural areas can be difficult to reach especially in a post-disaster setting, and poor people tend to have less social capital.

The empirical evidence across multiple developing country contexts generally finds the poorest have a lower ability to respond to weather shocks. For example, in response to the flooding and landslides in Nepal in 2011, only 6 percent of the very poor sought government support, compared with almost 90 percent of the better-off (Gentle *et al.*, 2014). In Thailand, the majority of government support after a flood benefited the well-off, with 500 baht per capita (around $14) going to the richest quartile, compared with 200 baht per capita (around $6) for the poorest quartile (Noy, Nguyen and Patel, 2021). A similar storyline is found in South India after the 2004 tsunami, with assistance biased away from the most disadvantaged and poorest castes (Aldrich, 2012). In addition, the main coping mechanisms such as selling livestock or other productive assets and reducing consumption can be harmful in the long run (Del Ninno, 2001; Heltberg, Oviedo and Talukdar, 2015; McCarthy *et al.*, 2018).

Even in cases where poor households receive post-disaster support, the amounts received are often too small to promote recovery. After the 1998 Great Flood in Bangladesh, 2/3 of poor households received transfers from the government, but on average the quantities accounted for only 4% of total household monthly expenditure (Del Ninno, 2001). Erman *et al.*, (2019) finds that this lack of access impedes recovery: after a flood in Dar es Salaam, Tanzania, higher-income households were able to recover much faster than poor households. However, one exception can be noted: the aid response to Cyclone Aila in 2009 in Bangladesh was well-targeted, and poor people were able to return to their pre-income level quicker than non-poor people (Akter and Mallick, 2013)

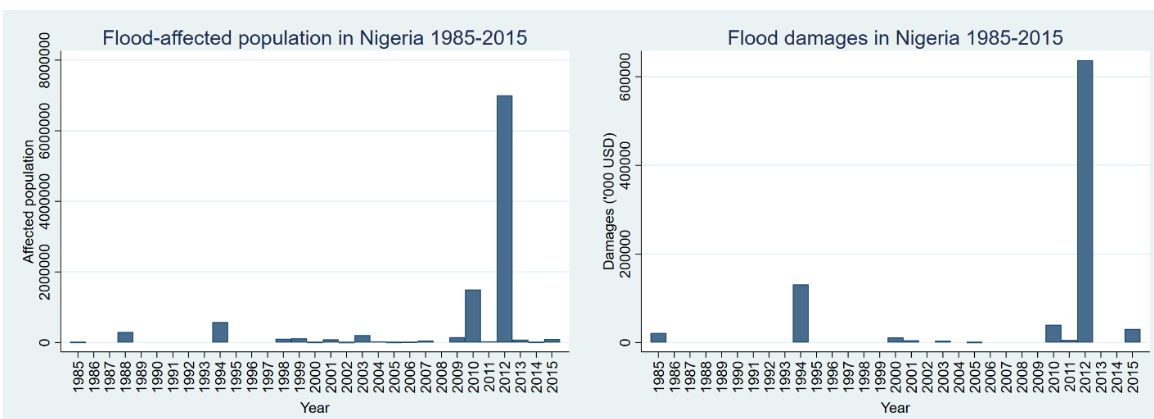
4.2.3. Floods in Nigeria and the 2012 event

Nigeria, often referred to as the “Giant of Africa”, is a large country located in West Africa that spans over 923,000 square kilometers and is home to over 200 million people, making it the continent’s most populous nation. The climate ranges from tropical rainforest in the south at the Niger Delta (where the Benue and Niger rivers drain into the Atlantic Ocean), tropical savannah which covers most of the country, and arid desert in the north and east. Despite being Africa’s largest economy, Nigeria is classified as a lower-middle income economy according to the World Bank, with a GDP per capita of $2,375 in constant 2010 USD as of 2019, with almost 40% of the population living below $1.90 per day (World Bank, 2021). The country has been urbanizing in recent decades, with 51% of the population living in urban areas as of 2019. Agriculture still plays a major role in the economy, accounting for 22% of GDP and 35% of total employment.

Floods are a regular occurrence during the rainy season each year (May to October), and have become the most severe natural hazard in recent years (Osumgborogwu and Chibo, 2017). In addition to heavy rainfall, flood events are often exacerbated by dam overflows and releases that impact households many miles downstream (Nzeribe, Nwokoye and Ezenekwe, 2014). Flooding affects millions across the country and is responsible for more displacement and property damage than any other hazard (Nzeribe, Nwokoye and Ezenekwe, 2014; Osumgborogwu and Chibo, 2017). Both global (Rentschler, Salhab and Jafino, 2022) and local (Nzeribe, Nwokoye and Ezenekwe, 2014) analyses find that around 18-20% of the total population are exposed to floods, with some studies suggesting over 70% are at risk in severe years (Agada and Nirupama, 2015). Recent trends find that flood exposure is rising: from 2000-2015, the share of population living in flood areas rose by 50% and climate change is likely to further increase exposure in the coming decades (Odjugo, 2009; Tellman *et al.*, 2020). In terms of how poverty and flood exposure interact in Nigeria, 15 million extremely poor people (below $1.90/day), 27 million moderately poor people (below $3.20/day), and 34 million poor people (below $5.50/day) live in high flood risk areas (Rentschler, Salhab and Jafino, 2022). Despite the importance of flooding across the country, comprehensive flood risk maps, a national flood risk management strategy, and safety nets to support households after an event remain inadequate (Olaore and Aja, 2014; Oladokun and Proverbs, 2016; Echendu, 2020).

Applying the framework on exposure, vulnerability, and ability to respond, Nigeria's socio-economic context, high agricultural reliance, low income per capita, lack of investment of flood infrastructure, and absence of a safety net suggests that any large-scale flood event is likely to have severe consequences, especially on the livelihoods of the poorest. This was evidenced during the unprecedented floods during the 2012 rainy season. Statistics compiled from The International Disasters Database collected by the Centre for Research on the Epidemiology of Disasters (EM-DAT) over the period 1985-2015 finds that in terms of affected population and economic damages, 2012 was an order of magnitude higher than recent trends (Figure 19) (CRED, 2023). The remainder of this paper focuses on this specific 2012 event which is likely to inform on the full impacts of floods across the country and provide implications for flood risk management in an era of increasing risk.

Figure 19. Population exposed to floods and economic damages in Nigeria, 1985-2015.



Source: Compiled from EM-DAT.

In March of 2012, months before the rainy season, the Nigerian Meteorological Agency predicted massive flooding; however, little preparation was done (NIMET, 2012; Agada and Nirupama, 2015). Consistent heavy rainfall occurred from July onwards, and by the end of September water reservoirs had overflown and water was released from dams across the country and in neighboring Cameroon, further exacerbating impacts (OCHA, 2012; Nzeribe, Nwokoye and Ezenekwe, 2014). The combination of heavy rainfall and water management decisions led to "unprecedented" floods which affected 30 of the country's 36 states, killed over 431 people, and displaced more than 1 million from their homes (Unah, 2017). A Post-Disaster Needs Assessment (PDNA) conducted by the Nigerian government (with support from international agencies) found that the event cost the economy $16.9 billion USD in economic damages and losses, or around 1.4% of country-level GDP (Federal Government of Nigeria, 2012). In terms of post-disaster aid, $180 million USD of federal and private funds were allocated to affected states by the end of 2012, accounting for only 1% of the total damage (OCHA, 2012; Unah, 2017). While the PDNA provided an overall figure to communicate the scale of the flood and identified which states and sectors experienced the largest economic impacts, these aggregate estimates may not tell the full story of the impact on people's livelihoods, especially the poorest. The economic impact figures were calculated by summing up damages (monetary replacement value of damaged durable assets) and losses (changes in the flows of goods and services) (Federal Government of Nigeria, 2012). The poorest households own or produce relatively little, which would not show up in the aggregate economic impacts; however, the livelihoods of these households are likely to be severely affected (Anttila-Hughes and Hsiang, 2013; Hallegatte, Vogt-Schilb, *et al.*, 2016; Botzen, Deschenes and Sanders, 2019).

Agada and Nirupama (2015) and Wizor and Week (2014) explore some of the local impacts of the 2012 event in Benue and Bayelsa states through household surveys. These studies uncover the origins, history, and problems that increased exposure and social vulnerability and reduced the ability for households to respond. In the cities of Makurdi and Otukpo in Benue, canals and drainage channels were blocked and filled up with waste, leading to higher exposure. Water levels rose to up to two stories, submerging homes and buildings completely and suffering irreparable damage, and leading to internal displacement. Agada and Nirupama (2015) also document that outbreaks of violence and competition for resources among ethnic, religious, and linguistic groups reduced the ability for the state to respond to the humanitarian crisis efficiently and equitably. In Yenagoa city of Benue state, Wizor and Week (2014) survey 465 households: around 50% report that settlements in the flood area and a lack of alternative livelihoods led to higher exposure to rainfall events. Between 70-85% of respondents also reported damage to school and health infrastructure as well as drinking water quality and sanitation facilities, leading to waterborne disease. In terms of coping strategies, 20% of households prepared mosquito nets and created embankments along the waterfront, while almost half were forced to relocate.

This paper makes three main contributions to the literature. First, it builds on the PDNA from the government and the two important local studies to provide a more comprehensive accounting of the impacts of Nigeria's 2012 flood on livelihoods. While Benue and Bayelsa states (which were previously studied) are located in the south and central parts of Nigeria, new research suggests that the northern part may be even more impacted by floods (Umar and Gray, 2022), and this paper explores impacts across the country. Furthermore, this paper provides quantitative evidence on the exposure, vulnerability, and ability to respond of households and includes an interview with an individual impacted in the village of Osomori. The second contribution of this paper is to use multiple modalities of measuring flood exposure to isolate the causal impact of the flood. I combine high-resolution remotely sensed data with household data on location and self-reported shocks to estimate who was directly impacted and create a plausible control group. Third, this paper explores how vulnerability is distributed across households. When examining vulnerability and the ability to respond, I employ micro-level panel data before and after the flood to identify impacts on agricultural yields, value, livestock, and consumption to better understand spillovers, interactions with local market effects, and how vulnerability relates to historical flood exposure, poverty, and other household characteristics.

4.3. Research Design

4.3.1. Research Questions

The focus of this paper is on the local impacts of Nigeria's 2012 floods on socioeconomic outcomes, with a focus on the exposure, vulnerability, and ability to respond of households. Motivated by the literature review presented above, I outline specific research questions below for each component of the risk framework (exposure, vulnerability, and ability to respond)[40].

Exposure:

[40] Certain research questions may fit into two categories – for example, consumption changes have been considered either in the "vulnerability" or "ability to respond" component of the framework in previous literature (Baez *et al.*, 2017; Hallegatte *et al.*, 2020). For this study, consumption changes are considered in the ability to respond component.

(1) Who was exposed to the 2012 event? How does exposure differ based on which flood metric is being used?
(2) Were poor households more or less exposed to the 2012 event compared to the average household?

Vulnerability:

(1) How did the flood impact crop production and value for affected farmers? Do impacts vary based on the type of crop?
(2) How were the impacts distributed across agricultural households? Were there any spillovers and local market effects?
(3) Did agricultural households that live outside historically flooded areas experience differential impacts?

Ability to respond:

(1) Did affected households change their consumption? Were consumption changes different for urban and rural households?
(2) How did households cope with the shock? Did affected farmers have to sell livestock? What was recovery like the year after?

4.3.2. Data

4.3.2.1. Measuring agricultural production, consumption, and food prices at the household level

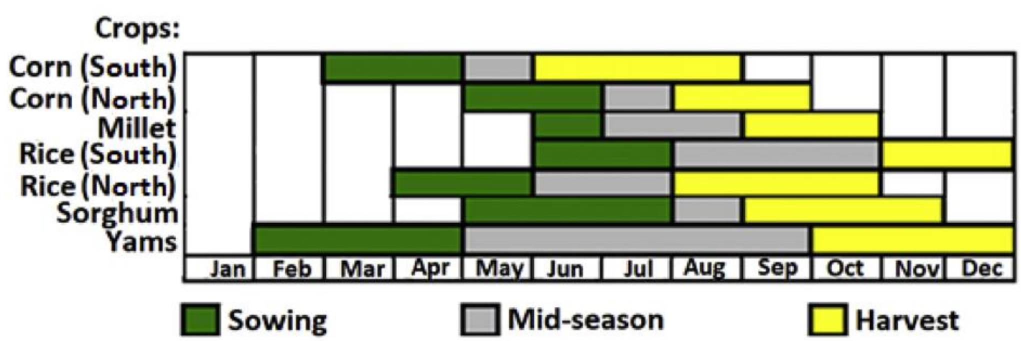
The dataset used to collect information on household agricultural activities, consumption patterns, and socio-economic characteristics comes from the Nigeria General Household Survey – Panel (GHS-Panel) [41]. This nationally-representative large-scale panel survey interviews the same 5,000 households (3,000 in rural areas and 2,000 in urban areas) across four waves: Wave 1 in 2010-11, Wave 2 in 2012-13, Wave 3 in 2015-16, and Wave 4 in 2018-19. The data for Waves 1 and 2 only are used in this paper for two reasons. First, given the interest in isolating the impacts of the 2012 flood event (which occurred just before Wave 2), including other waves of the survey may introduce

[41] This household survey is part of the World Bank's Living Standards Measurement Survey – Integrated Surveys on Agriculture, which is a coordinated set of publicly available panel surveys run in 8 countries in sub-Saharan Africa (Burkina Faso, Ethiopia, Malawi, Mali, Niger, Nigeria, Tanzania, and Uganda)

endogeneity by including unobservable factors that may influence the socio-economic outcomes of interest and be correlated with flood exposure (e.g. government policies, commodity price shocks, or other floods). Second, the agricultural variables of interest are not consistently measured in Waves 3 and 4, and the number of panel households tracked reduce by half. To reduce measurement error, these waves are currently excluded but plan to be used in future work[42].

Each wave of the survey includes two visits to the household and collects detailed information on agricultural practices, household socio-economic conditions, exposure to shocks, food price data, and the approximate geographic location of the household. In each wave, the household is visited twice: the first visit is termed "post-planting" (August-November) and the second is termed "post-harvest" (February-April of the following year). In each visit, surveys are conducted at the agricultural, household, and community level. The surveyors attempt to capture the seasonality of socio-economic conditions with this visitation schedule, however, it only roughly aligns with planting and harvest timeframes across Nigeria (Figure 20) (Shiru *et al.*, 2019). Furthermore, different parts of Nigeria experienced flood impacts at different times during the year: for example, the most severe impacts were felt by households in Bayelsa and Benue states in September and October.

Figure 20. Cropping calendar for selected crops in Nigeria.



Source: Shiru *et al.* (2019)

While the timing of the data collection does introduce measurement error across time and space, it does provide information before and after the flood for the same households. Additionally, it may be less prone to reporting bias since it is not a flood-specific survey (as is done for other studies, for

[42] More information on extensions of this paper using the Wave 3 and Wave 4 data is provided in the Conclusion section.

example: McCarthy *et al.* (2018)), but rather a general survey that occurs at a regular schedule and asks hundreds of questions on many topics. Another benefit of the survey is that the surveyors make their best attempts to track households that have moved, with an attrition rate of 9% across the 2 waves. For context, attrition rates for household surveys in developing countries range anywhere from 1-23% (Hagen-Zanker, Mallett and Slater, 2015).

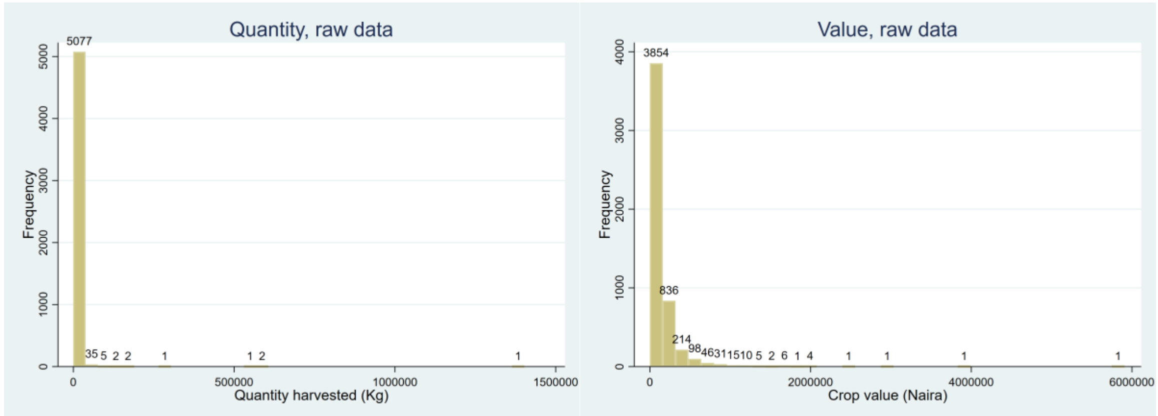
To measure agricultural outcomes for each farming household, I use two questions from the survey that measure output: one on quantity of crop production (measured in kg) and one on values of crop production (in Naira). As these two questions on production and value are only asked during the post-harvest visit of each Wave, I observe agricultural production and value at two points in time: 2011 (Wave 1) and 2013 (Wave 2). For quantity, the interviewer asks each agricultural household how many kilograms of each crop they harvested in the last season. I examine kg produced for all crops, and separately for the seven major crops grown among farmers in the sample which include maize, sorghum, cassava, beans/cowpea, millet, yams, and rice. While data on observed farm revenues as in other studies (Akter and Mallick, 2013) is unavailable in the GHS-Panel, the question on crop values is a proxy for potential revenues. Each farmer is asked how much they would receive (in Naira) if they were to sell all crops harvested this year (snapshots from survey questionnaire presented in Figure 21).

Figure 21. Snapshots from the survey on the two agricultural outcome variables of interest: production (left) and value (right)

6.		18.
How much did you harvest since the last interview? UNIT CODE KILOGRAMS (kg)...01 GRAMS (g)........02 LITRE (l)........03 CENTILITRE (cl)..04 MUDU.............05 OLODO............06 CONGO............07 PAINT RUBBER.....08 LARGE DERICA.....09 MEDIUM DERICA....10 SMALL DERICA.....11 MILK CUP.........12 CIGARETTE CUP....13 TIYA.............14 KOBIOWU..........15		If you had sold all [CROP] harvested since the last visit, what would be the total value?
QUANTITY	PROD. UNIT CODE	NAIRA

While these two questions on production and value are commonly used in the literature to measure agricultural outcomes (Del Ninno, 2001; Akter and Mallick, 2013; McCarthy *et al.*, 2018), they have several limitations. First, agricultural survey data in developing countries can be very noisy, often with outliers whose reported value is many standard deviations above the mean and does not appear realistic from a qualitative sense (Fraval *et al.*, 2019). The GHS-Panel data for Nigeria is no exception, as the distribution of the quantity and value outcomes are highly skewed (Figure 22). Following standard practice (BenYishay *et al.*, 2020; Meles, 2020), I winsorize the top-end of outliers at the 99^{th} and 95^{th} percentiles. This process retains all of the observations of the data, but assigns all values above the percentile to be equal to the percentile itself. For instance, 99^{th} percentile for the raw data on quantity is 28,300 Kg: for this threshold, all 30 observations above would be assigned 28,300 in the data. Another limitation is that I do not observe the date of the harvest for each household, which varies based on crop and region in Nigeria (Shiru *et al.*, 2019). If households were able to harvest before the flood hit their community, this may bias the results downwards. Additionally, the question on farmer value is a measure of potential rather than actualized revenue sales, which potentially introduces measurement error if farmers under- or over-estimate the value of their production and does not account for costs involved in accessing markets.

Figure 22. Distribution of raw data for crop production and crop value outcome variables.



Several factors may affect the relationship between exposure to floods and agricultural production, and I control for a number of these variables using data from the survey. Household structure may influence location decisions as well as labor supply available for agriculture, and is proxied using the number of adults in each household (McCarthy *et al.*, 2018). Additionally, households with some level of risk-sharing and market integration might be able to reduce losses from a flood as well as invest more into productive capital on the farm (Le Cotty, Maitre D'Hotel and Ndiaye, 2017), and I proxy for this using the presence of a bank account and distance to market[43]. Plot-level characteristics such as slope and the size of landholdings may influence how much farmland gets inundated when a flood hits, and is also correlated with the productivity and production capability of the land (Narloch, 2016; McCarthy *et al.*, 2018). In certain parts of sub-Saharan Africa, households may live far away from where their farmland is located (Masters *et al.*, 2013), but for households in the sample, farm location tends to be very close to the household location (<500m). As a result, I do not include a control for distance to farm from the household. Other household characteristics including age of the household head and elevation are also collected from the dataset.

The household survey data component of the GHS-Panel includes detailed information on each household's consumption at each visit, which provides 4 datapoints across the study period (2010, 2011, 2012, and 2013). Data is collected on food consumption (in the last 7 days), food expenditures (in the last month), non-durable consumption (in the last year), and on durable goods (in the last

[43] Asset ownership may also help with risk-sharing, and a previous version of this paper constructed an asset index for each household. The results of this paper are very similar with or without including this index as a control, with this asset data being omitted in the main analysis.

year), and is annualized and aggregated to provide the total consumption per capita for each household. While consumption is also commonly used in the literature to measure livelihood impacts of shocks (e.g. Baez *et al.*, 2017), it has its limitations. Importantly, examining consumption impacts alone may hide long-run livelihood impacts. If households sell or draw down their livestock in response to a flood, they may maintain their pre-shock level of consumption but lose a productive asset in the process (Heltberg, Oviedo and Talukdar, 2015). Second, like with the agricultural data, consumption information is self-reported and not observed. For example, wealthier households may under-report their consumption (Fraval *et al.*, 2019), which may introduce measurement error and introduce bias if household wealth is correlated with flood exposure.

I also collect data from the community component of the GHS-Panel to measure inflation-adjusted food prices at the local level. In the survey, there are approximately 10 households per community. While the community survey is not nationally representative (it covers 1/3 of all communities in Nigeria), the data does provide market prices for 321 communities for the staple crops: maize, sorghum, cassava, cowpea, millet, yams, and rice[44] in 2011 (Wave 1 Post-Harvest) and 2013 (Wave 2 Post-Harvest). Additional food price data at the state level is collected from the Nigeria Bureau of Statistics (Nigeria Bureau of Statistics, 2023). Data on livestock value and livestock sales is also collected to examine agricultural responses. In addition to quantitative data, I conducted a semi-structured interview with an individual whose relatives were directly impacted by the 2012 event to better understand how people experienced the event at the local level.

4.3.2.2. Data sources to measure flood exposure

Papers examining flood impacts across The Global South measure flood exposure either by asking people to self-report flood experience (Akter and Mallick, 2013; Opondo, 2013; Patankar, 2015) or by overlaying flood extents collected from satellites with an approximate GPS location of the household (Baez *et al.*, 2017; McCarthy *et al.*, 2018; Tiague, 2021). This study uses both metrics of measuring flood exposure in the same context.

[44] Similar to the agricultural data, the price data are highly skewed. Since there are no zero values (like there are in the agricultural data), I take the natural log of the price, and compare changes from Wave 1 to Wave 2 in the analysis.

The first exposure metric (henceforth: self-reported measure) relies on one module of the household survey that asks the household head to recall economic shocks experienced in the recent past. The exact question asked is *"Has your household been affected by [SHOCK] in the past 5 years?"* There are 22 different shock categories that each household provides a "Yes" or "No" response to, ranging from death of a household member, job loss, to weather shocks like floods for each of the previous 5 years before the survey interview (Figure 23). The main sub-question of interest is Shock Code 13 which asks if the household has experienced "Flooding that caused harvest failure"[45].

Figure 23. Shock questionnaire from household survey.

I'D LIKE TO ASK YOU ABOUT EVENTS THAT MAY HAVE AFFECTED YOUR HOUSEHOLD OVER THE LAST 5 YEARS.

SHOCK CODE	1		2	3				
	Has your household been affected by [SHOCK] in the past 5 years? YES..1 NO...2 (► NEXT SHOCK)		How many times has this occurred in the past 5 years?	In what years did this event occur?				
				2009	2010	2011	2012	2013
1	Death or disability of an adult working member of the household							
2	Death of someone who sends remittances to the household							
3	Illness of income earning member of the household							
4	Loss of an important contact							
5	Job loss							
6	Departure of income earning member of the household due to separation or divorce							
7	Departure of income earning member of the household due to marriage							
8	Nonfarm business failure							
9	Theft of crops, cash, livestock or other property							
10	Destruction of harvest by fire							
11	Dwelling damaged/demolished							
12	Poor rains that caused harvest failure							
13	Flooding that caused harvest failure							
14	Pest invasion that caused harvest failure or storage loss							
15	Loss of property due to fire or flood							
16	Loss of land							
17	Death of livestock due to illness							
18	Increase in price of inputs							
19	Fall in the price of output							
20	Increase in price of major food items consumed							
21	Kidnapping/Hijacking/robbery/assault							
22	Other (specify)							

The self-reported definition of exposure has the advantage of being as close to household's experience of the event as possible. However, these survey responses may not be accurate for several reasons. First, the exact wording of the survey asks if the household experienced a flood that *caused* a harvest failure. Specifics are not provided in the survey documentation as to what defines a "harvest failure", and each respondent's understanding of this term can vary widely.

[45] Shock Code 15 asks if the household experienced "Loss of property due to fire or flood". Since I cannot separate property damage from fire from property damage from flood, I did not use data from this sub-question.

Additionally, households who may have experienced a flood but not a harvest failure (for example, due to farm-level water management) or households who experienced large non-agricultural losses from the flood would not be identified as flooded using this survey question. For these reasons, when examining flood damages, I focus only on agricultural impacts in terms of crop production and value. In addition to answering “Yes” or “No” to each shock category, households are asked about their responses to the flood, and this data is also collected to qualitatively understand the ability for households to adapt.

Second, there is a long duration of recall in the survey questionnaire, which asks households to remember shocks which occurred in the past 5 years and may provide inconsistent results due to limitations in human memory (Fanta, Šálek and Sklenicka, 2019). In this case, the analysis for the 2012 flood in Nigeria may be less prone to measurement error. Households were asked about economic shocks in early 2013, which is close to when the floods occurred (July-October of the previous year). Seven times more households report being flooded in 2012 compared to the long-run average, suggesting this event to be extreme and likely to be remembered.

Third, the responses themselves may not be accurate. Households who did not experience a flood in 2012 may tell interviewers that they did experience a flood, and vice-versa. For example, if households believe that saying “Yes” to being exposed to the flood shock may be tied to support from either the government or an NGO, they may be more likely to report being flooded when they did not actually experience a flood (Erman *et al.*, 2018). However, in this context, I do not expect a bias in strategic reporting. Reporting bias is likely higher when the weather shock is more salient during the survey interview: many studies in the literature conduct an ex-post disaster-specific survey to examine the impacts of a flood or drought (Akter and Mallick, 2013; Patankar, 2015; McCarthy *et al.*, 2018). In these cases, investigating the impacts of the shock itself is the main reason for the survey and the salience to potential support is more prominent. In this paper, I use a large-scale nationally representative general household survey that has been going on since 2010, with multiple questionnaires per visit and over 15 modules per questionnaire. The questions of interest for this paper relating to the impacts of the 2012 flood are only a small component of the overall interview, which suggests reporting bias is unlikely.

Fourth, a household’s historical exposure to floods may affect their perceptions of whether they were flooded (Guiteras, Jina and Mobarak, 2015). To account for this potential bias, I complement the analysis by combining the approximate household location with satellite images on the flood extent and use risk maps to examine historical exposure to floods in Nigeria.

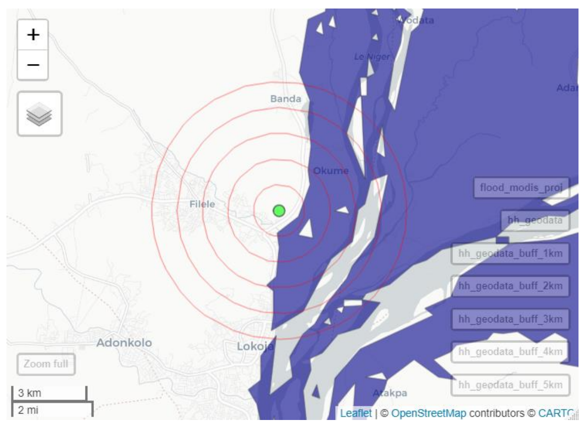
The second exposure metric (henceforth: satellite measure) combines satellite images on the observed flood extent during the study period with GPS information on the approximate household location from the GHS-Panel and may be less prone to human bias (Guiteras, Jina and Mobarak, 2015). Geographic information is collected in the survey which gives an estimate of the location of the household. Due to confidentiality reasons, the exact coordinate location of the households is not revealed; rather, the average of household coordinate locations in each survey cluster (primary sampling unit) is provided. There are around 10 households in each cluster. The centroid of these 10 household locations is calculated, and then offset by 0-2 kilometers in urban areas and 0-5 kilometers in rural areas, and is provided in the dataset in latitude-longitude format.

There are several options to measure the flood extent for historical events using satellite data. The most commonly used dataset in the literature is the Dartmouth Flood Observatory (DFO) (Kundzewicz, Pińskwar and Brakenridge, 2013; Chen, Giese and Chen, 2020; Kocornik-Mina *et al.*, 2020), which provides a global archive on large flood events (over 100 sq mi) that ranges back to the 1980s (Dartmouth Flood Observatory, 2018). The DFO's flood extent archives are derived from "news, governmental, instrumental, and remote sensing sources" (Dartmouth Flood Observatory, 2018), but precise details on how each of the sources are combined to produce a flood extent are unclear.

More recently, advances to more precisely measure flood extents comes from the NASA's Near Real-Time Global Flood Product (NRT), which was developed through a partnership with the DFO (Policelli *et al.*, 2017; NASA, 2023). The NRT is a daily, near-global, 250m resolution product showing flood water based on images collected by the Moderate Resolution Imaging Spectroradiometer (MODIS) instrument, which is based onboard NASA's Terra and Aqua satellites (NASA, 2022, 2023). While the product can detect many types of large-scale flooding, it is based on optical data and is unable to observe water under the ground during periods of cloud cover. While this is a limitation, cloud cover is not complete for many events, and shifts over a period of 1-2 days which reveals flood water below (Slayback, 2023). In the case of Nigeria's 2012 flood, the extent was calculated by the Global Facility for Disaster Risk Reduction (GFDRR) at the World Bank as part of the PDNA (Federal Government of Nigeria, 2012). It uses the NRT's 2-day product to account for cloud cover, and maps the flood extent during the study period (July-November) by aggregating each individual map (e.g. July 1-2) at the country level. This is the main satellite data used in this analysis to measure the flood extent of the 2012 event, and is referred to as MODIS-NRT in the rest of this paper[46].

Given the uncertainty regarding the exact household location, I calculate the intersection between flood and household at different buffer levels up to 5km which has previously been done in the literature (Winsemius *et al.*, 2018). The household location provided from the survey is buffered by 1km, 2km, 3km, 4km, and 5km, and is then intersected with the MODIS-NRT satellite map. An example of the buffer methodology is shown in Figure 24. In this example, the HH would be considered flooded (x=1) at the 1km, 2km, 3km, 4km, and 5km buffer but not flooded (x=0) without the buffer (at 0 buffer, there is no intersection with the flood map in blue). As the buffer levels increase, the resolution of the data becomes lower. According to the designers of the household survey, queries at low or medium resolution should be minimally affected by the imprecise coordinates of the households (Federal Republic of Nigeria and National Bureau of Statistics, 2016).

Figure 24. Buffering method used. Red circles represent the buffered boundaries of the HH coordinate, at 1km, 2km, 3km, 4km, and 5km. Blue area represents the MODIS-NRT flood map.



In addition to satellite data to capture the flood extent in 2012, I use flood hazard maps for Nigeria from the organization Fathom to better understand historical flood risk (Fathom, 2023). Flood hazard maps are modeled representations of risk and provide an indication of which areas are likely

[46] In the Appendix D1, a map is shown which compares the MODIS flood map to the Dartmouth flood map, and also includes the HH coordinate locations. The differences in scale are presented visually in this map, and for the reasons discussed above, the MODIS data is used to measure the 2012 flood in Nigeria for the remainder of this paper.

to be flooded for a specific event (e.g. 1-in-50 year flood, or 1-in-100 year flood). The data product used for Nigeria combines both pluvial (precipitation) and fluvial (riverine) flooding.

4.3.3. Identification Strategy

The first two research questions on exposure can be calculated by counting the number of households self-reporting flood, overlaying GPS information with the satellite flood extent, and separating statistics based on poverty status from the household survey. However, to examine the impacts of the flood on the vulnerability and the ability to adapt of households, a descriptive analysis at one point in time may not suffice. The empirical challenge in this context is to isolate the impact of the 2012 flood on crop production, crop value, consumption, and livestock sales from other variables that may influence the outcomes of interest and be correlated with flood exposure. Relatedly, households living in flooded areas may be systematically different than those living in unflooded areas. For example, as land prices may be cheaper in more flood prone areas, agricultural households with fewer resources may sort into flooded areas. These households may also be living on less productive land or have less ability to invest in their farm, which might drive agricultural outcomes and confound the impact of the flood.

This paper's identification strategy relies on two components to identify the impact of the flood on household outcomes using a difference-in-difference analysis. First, while the flood was not perfectly randomly assigned across the country, the 2012 event was unprecedented and widespread and can be considered a "quasi-random" shock. The review provided in Section 4.2.3 lends credence to this claim, as almost all states of the country were impacted, including areas historically affected by floods as well as "safer" areas. Figure 19 furthermore shows the scale of destruction in 2012 was an order of magnitude higher than any other floods in recent memory.

Second, this paper uses both flood measures introduced in Section 4.3.2.2 (self-reported and satellite metrics) to assign a treatment and plausibly-similar control group. The treatment group are all agricultural households who self-report a flood in 2012 from the household survey. As discussed in Section 4.3.2.2, many of the limitations of self-reported data are ameliorated in this case, and the self-reports are accepted at face value. Here, the challenge is to find a valid counterfactual for this treatment group. One plausible counterfactual are agricultural households who resided in the flood area in 2012 (using the satellite metric) but did not self-report a flood. These households in the control group are likely to face the same baseline exposure to floods as the treatment group, but due to the randomness of flooding events, they did not experience inundation on their plot of land.

There are two potential issues to this classification of treatment and control groups. First, survey responses in general may be prone to measurement error (Bertrand and Mullainathan, 2001). However, as reviewed in Section 4.3.2.2, the survey responses are less likely to be biased in this context due to the nature of the flood (minimizing recall issues), the agency implementing the survey (minimizing reporting bias), and the large-scale nature of the survey questionnaire (minimizing salience). Even if there is residual measurement error, it is unlikely to result in endogeneity if the error is not systematically correlated with the outcome variables of interest. This may be possible in this context due to the nature of the survey question on self-reported exposure which asks agricultural households to define flood exposure based on whether they experienced a flood that “caused harvest failure”. For example, a farmer may experience a small amount of crop loss (e.g. 10%), but still consider this a harvest failure and be considered as flood exposed in the analysis. The same farmer may also under-estimate their total production (or crop value), which may lead to a spurious relationship between flood exposure and agricultural outcomes. However, given that panel data is used, fixed effects can be employed to account for the “measurement ability” of each farmer and other characteristics that remain time-invariant.

The second concern may be that the self-reported flood exposure and satellite flood exposure are radically different measures. One might argue that the satellite data is an “objective” measure of flood exposure, where everyone within these flood zones are truly flooded, and the self-reported data is “subjective” and prone to the idiosyncrasies of household respondents (Guiteras, Jina and Mobarak, 2015). However, actual flood experience may not be homogenous within polygons defined as flooded by satellite data; in other words, within areas delineated by flood from space, a mottled mix of flood-affected and flood-unaffected households are contained. Even for the same event, different satellite products can identify very different areas affected by flood (Tellman, 2023), as is shown for Nigeria’s 2012 event in Appendix D1. Three different sources – MODIS, DFO, and Naharnet (from (Nzeribe, Nwokoye and Ezenekwe, 2014)) show very different areas affected by the same event.

Flood impacts vary greatly even at very local scales, as geographical determinants and water dynamics can inundate certain areas but leave neighboring homes unaffected (Patankar, 2015; Hallegatte, Bangalore, *et al.*, 2016; Winsemius *et al.*, 2018). This is also the case for Nigeria’s 2012 event. The household survey contains a community questionnaire that is administered to elders in each local government authority (sub-district) and includes questions on shocks experienced. Even within communities reporting a flood in 2012, some households within the community report a flood at the household-level while many others do not[47]. Additionally, evidence from a qualitative interview conducted with a key informant finds heterogenous flood exposure within one village (Osomori), with some areas totally inundated (in upland areas) while neighboring lands were spared. Overall, due the unprecedented nature of the flood, the low likelihood of measurement error of self-reported data, and the heterogeneity within satellite areas, the control group may be a plausible counter-factual for treated households.

The regression specification is presented below in Equation 4, where $\mathbf{Y_{it}}$ is the outcome of interest (production, value, consumption, livestock) for household i in survey wave t, $\boldsymbol{\gamma_i}$ represents household-fixed effects, $\boldsymbol{\delta_t}$ the time-fixed effects, and $\mathbf{D_{it}}$ indicates whether the household was flooded in 2012. Households who self-report a flood in 2012 are in the treatment group ($\mathbf{D_{it}} = \mathbf{1}$) while households living in the satellite flood area who did not report a flood in 2012 are considered as the control group ($\mathbf{D_{it}} = \mathbf{0}$). No controls are added in the main specification, but are included in robustness checks presented in Section 4.4.4. Household-fixed effects are included to control for unobserved factors that do not change over time (e.g. soil quality) and time-fixed effects are included to account for external characteristics occurring during Wave 1 (2011) and Wave 2 (2013) which include variables such as elections or government policies. The analysis compares outcomes at two points in time before (post-harvest Wave 1 conducted in Feb-April 2011) and after the flood (post-harvest Wave 2 conducted in Feb-April 2013)[48].

Equation 4. Regression specification to identify the impact of the 2012 floods on agricultural households

$$Y_{it} = \gamma_i + \delta_t + \beta D_{it} + \varepsilon_{it}$$

To supplement the quantitative analysis, a semi-structured interview was conducted with an individual whose sister and immediate family were impacted by the flood in 2012. At the time of

---

[47] This can also be seen visually in Figure 26.

[48] While it is possible farmers who were severely affected by the flood had to migrate and dropped out of the sample, flood exposure is not strongly correlated with missing households (0.07).

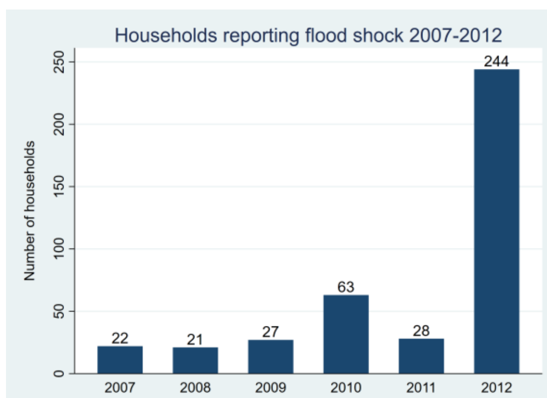
the flood, the interviewee's family lived in the village of Osomori in Anambra state, residing on the banks of the Niger river in the southern part of the country.

4.4. Results

4.4.1. Exposure

Annual self-reported exposure results from 2007-2012 are presented in Figure 25 for the panel of 2,563 agricultural households interviewed in Wave 1 and Wave 2. From 2007-2009, fewer than 1 percent (around 25 households in the survey) report flooding that caused harvest failure, this figure rises to 2.5% in 2010 (63 households), and drops back to 1% in 2011. In 2012, the share dramatically increases, with almost 10% of households self-reporting a flood that caused a harvest failure, with 244 "Yes" responses[49]. These descriptive findings provide empirical support to the qualitative evidence presented in the literature review suggesting the floods in 2012 were a uniquely extreme event.

Figure 25. Number of households reporting a "Flood that caused harvest failure" from 2007-2012 (of the panel of 2,563 households).



For the satellite metric, the number of agricultural households categorized as flooded depends on the size of the buffer used. Without buffering the household coordinate, only 38 households are in the MODIS-NRT satellite flood area. As the buffer gets larger, more households are classified as

[49] The data for 2013 shows a very small number of households reporting a flood shock that year, but since the data was collected before the rainy season, it is omitted from the figure on self-reported shocks over time. Data from Wave 3 of the survey confirm that the year 2013 had few self-reports of floods, similar to the trend from 2007-2011.

flooded, with the number rising to 207 at 3km and 394 at 5km as presented in Table 19. These two buffers are selected to account for the uncertainty in the household location (Winsemius *et al.*, 2018). These figures suggest between 8-15% of agricultural households were exposed to the flood, which is in line with the 10% figure from the self-reported data as well as estimates from the literature: Nzeribe, Nwokoye and Ezenekwe (2014) find that in local government areas that were affected by the flood, 12% of the total population was exposed in 2012.

Table 19. Number of households exposed to the 2012 floods calculated by intersecting the MODIS satellite with the HH location (at different buffers)

Method	Number exposed
MODIS (no buffer)	38
MODIS (HH location buffered 1km)	102
MODIS (HH location buffered 2km)	151
MODIS (HH location buffered 3km)	207
MODIS (HH location buffered 4km)	332
MODIS (HH location buffered 5km)	394

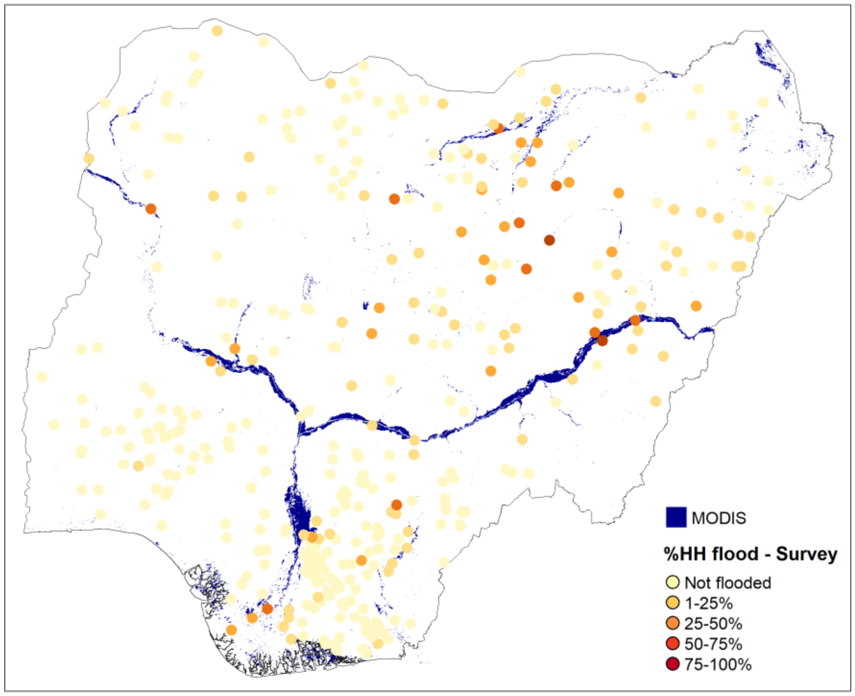
Are poor agricultural households more exposed to the 2012 flood compared to non-poor households? By combining these estimates on exposure with data on consumption, I calculate the share of flood-exposed households in the bottom 20% of the consumption distribution compared to all households, following Winsemius *et al.* (2018). Using this definition of poverty, 10.3% of households in the poorest quintile self-report being flooded, compared to 9.5% of the total sample. For the satellite metric at 3km, 10.1% of households in the poorest quintile reside in flood zones compared to 8.1% of all households. Using the 5km buffer, 15.6% of poor households are exposed compared to 15.4% for all households. For both self-reports and satellite exposure using the 5km buffer, it does not appear that poor households are over-represented in the flood areas. However, for the satellite exposure using the 3km buffer, poor households do appear to be over-represented in flood areas compared to the total population.

These results lie in the middle of studies exploring the poverty-flood exposure relationship in Nigeria. In a global modelling study of 52 countries, Winsemius et al. (2018) find that poor people (measured using a wealth index) are 52% more likely to be flooded in Nigeria, compared to the general population. The results of another global study using more detailed flood data and a consumption definition of poverty find that in Nigeria, poor people are 12-60% less likely to be exposed to floods, depending on the poverty line used (Rentschler, Salhab and Jafino, 2022). The differences found across studies could be due to the different definitions of poverty (asset vs.

consumption, or thresholds used), samples studied (Winsemius *et al.* (2018) and Rentschler, Salhab and Jafino (2022) include rural and urban households), as well as the different flood maps used. This paper uses the historical flood extent from 2012 from NRT-MODIS, Winsemius *et al.* (2018) use a global dataset of riverine flood risk called GLOFRIS, while Rentschler, Salhab and Jafino (2022) employ a higher resolution global flood dataset from Fathom.

The results on flood exposure from this study also presented spatially in Figure 26. The dots colored yellow to red represent the self-reported exposure metric: that is, share of households within each cluster that report experiencing flooding that caused harvest failure in 2012. There appears to be higher concentrations of flood reports in the south west and north east of the country, but self-reports to the 2012 event seem to be widespread in the survey data. The satellite data, alternatively, appears to represent overflow associated with the two main rivers of the country (Niger and Benue) and their confluence in the south west. There is some flood inundation in the north and north east, likely representing flood from smaller rivers. The self-reports and satellite are positively correlated, but only slightly, with a correlation coefficient of 0.20 (5km) – 0.23 (3km). While this low correlation may appear surprising, these two metrics differ based on the source and may measure different aspects of the flood (with the satellite likely capturing large river overflow). In general, the flood exposure metric from the self-reports can be interpreted as an individual-level metric, while the satellite exposure metric can be inferred as an area-based indicator of flood.

Figure 26. Exposure to the 2012 flood measured through self-reports (dots) and overlay between geographical location and MODIS-NRT satellite data.



Note: The “share of HH flooded” presents information on how many of the 10 households within each survey cluster report experiencing a flood that caused harvest failure.

4.4.2. Vulnerability

While the descriptive results on exposure explored in the previous section provide an understanding of who was exposed across the entire country, examining the causal impact of the flood on crop production, crop value, consumption, and livestock impacts use the identification strategy outlined in Section **Error! Reference source not found.**. For the two following sub-sections o n vulnerability (crop production and value) and the ability to adapt (consumption and livestock impacts), households who self-reported a flood are considered treated and households who live in the satellite flood zone but did not self-report a flood in 2012 are in the control group. When selecting the buffer size, I use the 3km buffer threshold which accounts for the uncertainty in the precise household location, but also provides a match which is geographically more similar to the treated group. In contrast, the 5km buffer threshold may capture households in semi-urban or peri-urban areas, who may be very different from agricultural households who self-report a flood (Federal Republic of Nigeria and National Bureau of Statistics, 2016). This provides a sample of 244 households in the treated group and 140 households in the control group, when using the 3km buffer.

4.4.2.1. *Descriptive statistics*

Table 20 shows that treatment (flood-affected) and control (within satellite flood zone, but not directly affected) households are broadly similar before treatment in terms of household demographics, wealth, risk coping strategies, and location characteristics. Following (McCarthy *et al.*, 2018), I proxy for these socio-economic characteristics by examining the number of adults, the age of the household head, bank account presence, poverty status[50], landholdings, distance to market, and the slope and elevation of the agricultural plot. The data is top-end windsorized at the 95th percentile to account for outliers following BenYishay *et al.* (2020). Treatment and control households display a high similarity in terms of number of adults, bank account presence, poverty status, landholdings, slope, and elevation. Control households show an older household head than treatment households, although this difference is only marginally significant. In terms of distance to market, treated households are around 12km further from the nearest market, with this difference being statistically significant at the 1% level.

Table 20. Household characteristics before the flood for treatment and control groups.

Characteristics	Treatment	Control	T-stat (diff)	P-value
No. of adults	3.67	3.72	0.25	0.80
HH head age	48.18	51.06	**1.91**	**0.06**
Bank account	0.13	0.14	0.44	0.66
Poverty status	0.46	0.43	-0.58	0.56
Landholdings	2020.34	2020.44	0.00	1.00
Distance mkt	75.77	63.27	**-3.12**	**0.00**
Slope (%)	2.53	2.20	-1.41	0.16
Elevation (m)	320.96	286.57	-1.25	0.21
No. HH	244	140		

Note: Landholdings and the distance to market are represented in square kilometres.

4.4.2.2. *Impacts on crop production and value*

In terms of crop production, I find that farmers who were flooded in 2012 lost around 21% of total production, with heterogeneity by crop. Full results are presented in Table 21 for aggregate crop

[50] A poverty line of 49,994 Naira obtained from the World Bank (inflation-adjusted) allows to separate poor from non-poor households (with households coded as being poor if their consumption in the year 2011 or 2013 falls below this threshold).

production impacts and broken down for each of the 7 major crops planted in Nigeria (maize, sorghum, cassava, bean, yams, millet, and rice). Agricultural households flooded in 2012 experienced a statistically significant decline in total crop production by around 635kg, or 21% of the mean before the flood. This estimate is consistent but slightly lower than the 30% drop found after floods in Malawi (McCarthy *et al.*, 2021) and Tanzania (Tiague, 2021).

Crop-specific results find large, significant, and consistent drops in sorghum, cassava, and millet but find other crops – including maize, beans, yams, and rice – to be unaffected. Crop-specific impacts from this analysis are inconsistent with previous studies at the state-level which found that rice production decreased the most by 31%, followed by a 17% drop in yam yields, 14% for cassava, and 8% for maize, and 5% for sorghum (Federal Government of Nigeria, 2012). There are two potential explanations to reconcile the differing results. First, the results presented in this analysis compare outcomes before and after for broadly similar households. The results at the state-level are for what share of planted land was unharvested due to the flood, and do not compare production before and after the flood. Second, the results in the household survey data do not account for when crops are harvested across different parts of the country in relation to the flood timing and may explain the inconsistent results. The results for crop value are presented in Table 22. Households directly affected by floods experience a large and statistically significant drop in crop value by around 70,000 Naira (~435 USD)[51]. This is a substantial decline for affected farmers, and represents a drop of 53% compared to the mean value before the flood.

Table 21. Impact of flood on agricultural production, for all production and the major crops.

	(1) All Crops	(2) Maize	(3) Sorghum	(4) Cassava	(5) Bean	(6) Yams	(7) Mill	(8) Rice
Flood	-635.0*	14.7	-343.7***	-398.9**	-23.8	515.5	-158.4*	-32.7
	(355.5)	(107.0)	(105.9)	(197.3)	(41.3)	(591.7)	(92.4)	(73.8)
Obs	768	376	416	182	358	162	290	144
N of HH	384	188	208	91	179	81	145	72
Controls	No	No	No	No	No	No	No	No
HH FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Cluster	HHID	HHID	HHID	HHID	HHID	HHID	HHID	HHID

[51] Importantly, the empirical patterns in the data that show large drops in crop production, and especially in crop value, are a form of verification of the self-reported data, which increases confidence in the empirical set-up and identification strategy.

Outliers	W95	W95	W95	W95	W95	W95	W95	W95

Note: Results include household fixed effects and time fixed effects and are clustered at the household level. W95 indicates the dependant variable is top-end windsorized at the 95th percentile. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 22. Impacts of the flood on crop value for affected farmers.

	(1) Crop Value
Flood	-68,599***
	(12,822)
Obs	768
N of HH	384
Controls	No
HH FE	Yes
Time FE	Yes
Cluster	HHID
Outliers	W95

Note: Crop value metric is measured in Naira. Results include household fixed effects and time fixed effects and are clustered at the household level. W95 indicates the dependant variable is top-end windsorized at the 95th percentile. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

4.4.2.3. Spillovers and local market impacts

While most studies in the literature either use self-reported data or satellite data to understand flood impacts on households, I combine the two measures to better understand spillovers in relation to Nigeria's 2012 flood. The two flood measures in this paper can be thought of at different scales: the self-reports are an estimate of individual-level exposure, while the satellite metric can be informative on the level of flooding in a particular area. An agricultural household's impact from a flood may be determined by both its own individual-level exposure but also that of their neighbors (e.g. the surrounding area) and are explored in this section.

To tease out potential spillover impacts, I explore impacts on agricultural value for households within flood areas (satellite) who also self-report being flooded compared to households who live in the flood area but do not report being flooded. To do so, I run the same regression framework as Equation 4 but the treatment variable is replaced with an interaction of the satellite and self-reported measures. The results are presented in Table 23. Column 1 shows that households who live in a flooded area and individually report a flood experience a drop in value of 41,000 Naira, which is statistically significant at the 1% level. Column 3 provides estimates of the effect of living in a flooded area on crop value, if the household does not itself report an individual-level exposure.

Conversely for this sample, I find that crop values increased for this subset, which is statistically significant at the 1% level and economically meaningful. These individually-unaffected households may have benefitted from living in a flooded area, with an increase in crop value of around 28,000 Naira (~$175 USD), or around 30% compared to the mean.

Table 23. Effect of living in a flooded area on crop value, for households who also report being individually affected (Column 2) versus those who do not report being individually affected (Column 3).

	(1) Yes individually affected	(2) No individually affected
Flood (Satellite)	-41,186*** (13,732)	28,453*** (10,144)
Obs	768	768
N of HH	384	384
Controls	No	No
HH FE	Yes	Yes
Time FE	No	No
Cluster	HHID	HHID
Outliers	W95	W95

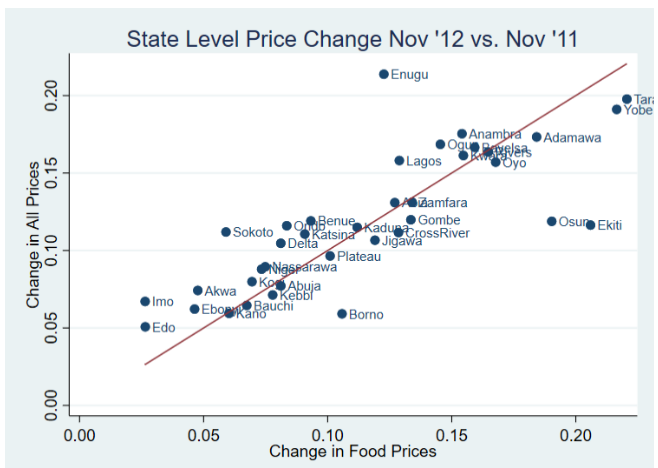
Note: Crop value metric is measured in Naira. Results include household fixed effects and time fixed effects and are clustered at the household level. W95 indicates the dependant variable is top-end windsorized at the 95th percentile. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Time fixed effects removed due to collinearity.

Why might households living in affected areas benefit from an extreme event like the 2012 floods? One hypothesis is that the price of food increases after a local production shock (Barrett, 2010). For example, Baez et al. (2017) documents a 17% increase in food prices after Tropical Storm Agatha hit Guatemala in 2010. If flooded areas experience an increase in food prices, households who have retained their harvest may be able to benefit from higher sales prices at the market.

I test this hypothesis in two ways, by using state-level data on food price changes from the Nigeria Bureau of Statistics as well as market price information collected as part of the household survey. I collected data at the state level to measure the change in all prices (e.g. general inflation) and the change in food prices comparing before the flood (in November 2011) to after the flood (in November 2012). I might expect areas that were hardest hit by the flood (e.g. Benue and Bayelsa) to experience more inflation in food prices compared to general inflation. The results are presented in the scatter-plot in Figure 27. Qualitatively, I do not find that areas with higher levels of flood exposure in 2012 experienced a larger increase in food prices compared to general inflation. Most estimates fall close to the 45 degree line (in red), suggesting food prices increases were similar to price increases for all goods comparing post-flood to pre-flood.

I further investigate the state-level price changes by categorizing each state as being affected or unaffected by the flood based on economic damage estimates from the Government of Nigeria (Federal Government of Nigeria, 2012). Twelve states (Adamawa, Anambra, Bayelsa, Benue, Delta, Edo, Jigawa, Kebbi, Kogi, Nassarawa, Rivers, and Taraba) experienced considerable flood damages while the other 25 did not. A t-test without controls comparing the change in food price (vs. general inflation) finds no differences, with the flooded sample experiencing a change of -.006 and the unflooded sample of -.002, with a t-statistic of 0.29. Qualitatively and quantitively there does not appear to be any change in food prices for flood-affected areas at the state level.

Figure 27. Scatter-plot at the state level to compare the change in food prices (x-axis) to the change in all prices (y-axis).



Note: A 45 degree line (in red) indicates food prices and all prices increase at the same rate. Data is collected from the Nigeria Bureau of Statistics.

However, data at the state level may obscure local impacts in certain communities and may not be comparing similar households, some of whom were affected by the flood and others who were spared. To complement the state-level analysis, I collect food price data from the community interview as part of the same household survey. The data provides market prices for all staple crops (maize, sorghum, cassava, cowpea, millet, yams, and rice) in 321 communities that were part of the survey (10 households are surveyed in each community). This only covers 1/3 of all communities in Nigeria and is not nationally representative but can provide an indication of the market prices that households in the survey face[52]. The price data varies greatly across less expensive and more expensive crops. As I am examining the percentage change in price, I log transform the price variable as it is more analytically tractable and more easily interpretable.

I classify each of these 321 communities into flooded or not flooded using the self-reported method, with communities classified as flooded if at least one household within the community individually reported a flood in 2012[53]. The results are presented in Table 24 using the same regression framework with time and community fixed effects. Communities that were flooded experienced an increase in log food prices by .302, which is statistically significant at the 5% level. This translates to a 35% increase in local prices (measured by Naira/kg), which would explain why households who live in a flooded area but who were not individually impacted experienced a 20% increase in crop value.

Table 24. Change in log food prices before and after the flood, for communities impacted using the self-report definition.

	(1) Log Food Price
Flood (Self-reported)	.302** (.147)
Obs	642
N of communities	321
Controls	No
Community FE	Yes
Time FE	Yes
Cluster	Community

Note: Prices include all major crops: maize, sorghum, cassava, cowpea, millet, yams, and rice. Community fixed effects and time fixed effects are included and results are clustered at the community level. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

[52] In this case, it is unlikely that sampling of the communities was correlated with price changes, which is likely to reduce concerns of biased estimates.
[53] The first-choice analysis would have followed a similar regression framework from the results presented on crop production and value, but a shapefile of communities was unavailable, and I was unable to link that data to the satellite flood exposure. Future research to collect this spatial data can isolate communities that were directly affected (based on self-reports) versus those located in flood areas but not directly affected.

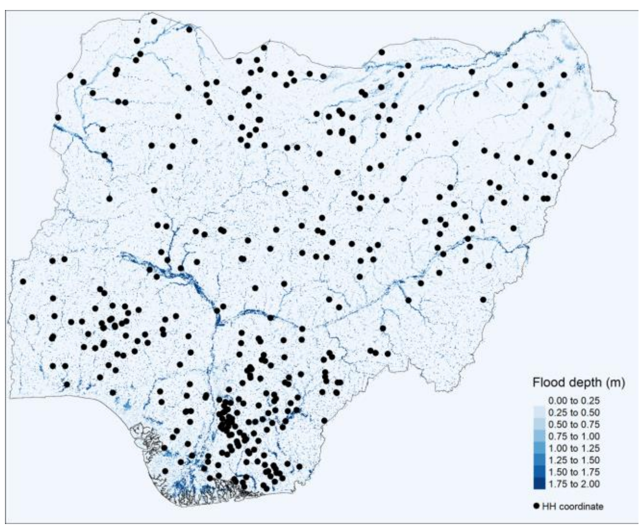
4.4.2.4. Heterogenous impacts

4.4.2.4.1. Historical exposure to floods

The distribution of agricultural consequences of the flood depend not only on impacts across space but also across time. Households who experience flood events regularly might be better adapted to extreme years like 2012 and thus be impacted less (Guiteras, Jina and Mobarak, 2015). This section explores heterogenous effects on agricultural production and value for households inside and outside of historical flood zones.

To measure historical flood exposure, I collect data on the expected flood extent and depth across Nigeria for a modeled event similar in scale to the 2012 event from the flood consultancy Fathom (Fathom, 2023). I combine this flood hazard map with the approximate household locations from Wave 1 of the survey, to separate households who live in historically flooded areas from those who live in historically unflooded areas. The overlay is presented in Figure 28. I run the same regressions framework from Equation 4 separately for households inside the historical flood zones and for those outside of historical flood zones.

Figure 28. Fathom flood risk map (modelled flood extent and depth for a 50 year return period event), overlayed with the approximate household locations from the survey.



Households who live outside of the historical flood zone (e.g. those areas not in the Fathom flood map) experience more significant impacts on crop production and value (Table 25). For crop production, households outside of historical flood zones experienced crop production losses that are twice as large as those in historically flooded zones (Columns 1 and 2). While both point estimates are not statistically significant, the coefficient for households outside of flood zones (Column 2) approaches significance at the 10% level. For crop value, a similar pattern emerges of households outside of historically flood zones experiencing larger impacts. Those inside historical flood zones experience a drop of 57,000 Naira compared to 84,000 Naira for households outside historical flood zones (Columns 3 and 4). Both estimates are statistically significant at 1% level. The results suggest that households living outside of historical flood zones lost around 2/3 of crop value due to the flood. Taking the production and value estimates together, these results provide strong evidence that households who less frequently experienced floods may have been more severely impacted by the 2012 event. One potential explanation is that households commonly exposed to floods may be more adapted and take protective measures to reduce farm-level impacts (Guiteras, Jina and Mobarak, 2015) compared to households in historically unflooded areas.

Table 25. Heterogenous impacts on crop production and value for households historically inside of flood zones (Columns 1 and 3) versus households historically outside of flood zones (Columns 2 and 4).

	(1) In historical zone – crop production	(2) Out historical zone – crop production	(3) In historical zone – crop value	(4) Out historical zone – crop value
Flood	-409.1	-836.7	-57,483***	-84,341***
	(478.9)	(519.2)	(17,282)	(18,545)
Obs	440	328	440	328
N of HH	220	164	220	164
Controls	No	No	No	No
HH FE	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes
Cluster	HHID	HHID	HHID	HHID
Outliers	W95	W95	W95	W95

Note: Households are split into those residing inside and outside of historical flood zones using the same regression framework from Equation 4. Results include household fixed effects and time fixed effects and are clustered at the household level. W95 indicates the dependant variable is top-end windsorized at the 95th percentile. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

4.4.2.4.2. Heterogenous impacts across household characteristics

In addition to historical flood exposure, household characteristics may interact with flood exposure in meaningful ways and uncover heterogenous impacts in terms of crop production and value. In particular, impacts may differ based on bank account presence, size of landholdings, distance to market, and steepness of slopes of the agricultural plot=(McCarthy *et al.*, 2018; Narloch and

Bangalore, 2018; Hallegatte *et al.*, 2020). This section examines these characteristics by interacting flood exposure with household characteristics using the regression framework presented in Equation 4.

Table 26, Table 27, Table 28, and Table 29 present the results for bank account, size of landholdings, distance to market, and slope. For crop production, households who are affected by flood and do not have a bank account experience a strong reduction of around 700kg which is significant at the 10% level, compared to no change for households with a bank account (Columns 1 and 2 of Table 26). In terms of crop value, both categories of households experience a large decline of around 60,000-70,000 Naira which is statistically significant at the 1% level.

Table 27 displays the results for the size of landholdings. For crop production, there is no difference between households with small versus large landholdings. In both cases, production drops by around 550kg but is not statistically significant. In terms of crop value, while declines are statistically significant for both categories, households with smaller landholdings report experiencing a larger drop of around 70,000 Naira which is twice the size for households with large landholdings (Columns 3 and 4). This may reflect the inability of farmers with small landholdings to diversify in response to a shock; if a flood enters their plot, a larger share is more likely to be impacted compared to farmers with larger landholdings.

Table 28 presents results based on the distance to market. In terms of crop production, there does not appear to be any difference based on distance to market, with both categories of households experiencing a drop of 650-700kg that is approaching statistical significance. In terms of crop value, households further from market experience larger declines of around 80,000 Naira (Column 4) compared to around 60,000 Naira for households closer to the market (Column 3). This greater loss in crop value for households further from markets suggests that in remote locations, production shocks are less likely to be replaced with imports. Moreover, the higher value of loss in remote locations may also reflect the higher food prices for staple goods.

Table 29 shows heterogeneous impacts based on the steepness of slopes. Households on less steep slopes report more losses in terms of crop production and value. The difference is particularly prominent for crop value (Columns 3 and 4), as households with steeper slopes experience no change, while those living on less steep slopes display a decrease of around 70,000 Naira. This may be a result of increased waterlogging in these locations, as floods may concentrate in areas with flatter slopes compared to locations with steeper slopes. With more waterlogging (for longer periods of time), crops may be impacted to a larger degree compared to agricultural households living on steeper slopes where floodwaters may pass through more quickly.

Table 26. Heterogenous impacts in terms of crop production and crop value based on bank account presence

	(1) Yes bank account – crop production	(2) No bank account – crop production	(3) Yes bank account – crop value	(4) No bank account – crop value
Flood	29.4	-693.6*	-78,682***	-60,570***
(Self-reported)	(838.6)	(365.6)	(28,932)	(12,835)
Obs	768	768	768	768
N of HH	384	384	384	384
Controls	No	No	No	No
HH FE	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes
Cluster	HHID	HHID	HHID	HHID
Outliers	W95	W95	W95	W95

Note: Results include household fixed effects and time fixed effects and are clustered at the household level. W95 indicates the dependant variable is top-end windsorized at the 95th percentile. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 27. Heterogenous impacts in terms of crop production and crop value based on size of landholdings

	(1) Large landholdings – crop production	(2) Small landholdings – crop production	(3) Large landholdings – crop value	(4) Small landholdings – crop value
Flood	-526.3	-595.2	-34,676*	-73,432***
(Self-reported)	(474.6)	(421.5)	(18,720)	(14,624)
Obs	768	768	768	768
N of HH	384	384	384	384
Controls	No	No	No	No
HH FE	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes
Cluster	HHID	HHID	HHID	HHID
Outliers	W95	W95	W95	W95

Notes: Large landholdings are defined as households with holdings above the mean of 1,898 sq m and those with small landholdings are defined as those below the mean of 1,898 sq m. Results include household fixed effects and time fixed effects and are clustered at the household level. W95 indicates the dependant variable is top-end windsorized at the 95th percentile. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 28. Heterogenous impacts in terms of crop production and crop value based on distance

to market

	(1) Long mkt distance – crop production	(2) Short mkt distance – crop production	(3) Long mkt distance – crop value	(4) Short mkt distance – crop value
Flood	-648.6	-683.9	-82,872***	-57,775***
(Self-reported)	(535.6)	(416.1)	(18,445)	(14,137)
Obs	768	768	768	768
N of HH	384	384	384	384
Controls	No	No	No	No
HH FE	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes
Cluster	HHID	HHID	HHID	HHID
Outliers	W95	W95	W95	W95

Notes: Households with a long distance to market are defined as those above the mean of 71.5 km, while those with short distance to market are those below the mean of 71.5 km. Results include household fixed effects and time fixed effects and are clustered at the household level. W95 indicates the dependant variable is top-end windsorized at the 95th percentile. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 29. Heterogenous impacts in terms of crop production and crop value based on steepness of slopes

	(1) More steep slope – crop production	(2) Less steep slope – crop production	(3) More steep slope – crop value	(4) Less steep slope – crop value
Flood	-848.3	-798.0**	-6,553	-68,467***
(Self-reported)	(883.8)	(386.2)	(35,825)	(14,501)
Obs	768	768	768	768
N of HH	384	384	384	384
Controls	No	No	No	No
HH FE	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes
Cluster	HHID	HHID	HHID	HHID
Outliers	W95	W95	W95	W95

Notes: Households with steeper slopes are those above the mean of 3.18% while those with less steep slopes are below the mean of 3.18%. Results include household fixed effects and time fixed effects and are clustered at the household level. W95 indicates the dependant variable is top-end windsorized at the 95th percentile. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

4.4.3. Ability to respond

In addition to understanding who was exposed to the event and how agricultural households were immediately impacted, this section aims to understand the ability for households to respond to the shock. Here, I conduct quantitative analyses to uncover consumption impacts for both agricultural and urban households, livestock impacts for agricultural households, and present evidence from an interview to understand how affected households responded, with a focus on the village of Osomori in Anambra state.

4.4.3.1. Consumption and livestock impacts, and potential poverty traps

Prior research on weather shocks across The Global South finds heterogenous responses in household consumption: sometimes consumption goes down (Rodriguez-Oreggia *et al.*, 2013; Baez *et al.*, 2017) and in other cases consumption remains unchanged (Baez, Caruso and Niu, 2020; McCarthy *et al.*, 2021). Here, I explore total consumption and food consumption as the main outcomes of interest. The results presented in Table 30 find that for both flood measures, no changes in total or food consumption per capita are detected.

Table 30. Impact of the flood on total consumption per capita for all households.

VARIABLES	(1) Total Consumption	(2) Food Consumption
Flood	-1,344	-1,206
	(5,928)	(4,372)
Observations	768	768
Number of hhid	384	384
Controls	No	No
Household FE	Yes	Yes
Time FE	Yes	Yes
Cluster	HHID	HHID

Notes: Total consumption is measured in Naira. Results include household fixed effects and time fixed effects and are clustered at the household level. W95 indicates the dependant variable is top-end windsorized at the 95th percentile. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

What might explain the lack of change in consumption? In response to a major shock like a flood or a drought, households close to the subsistence level may be able to smooth consumption in the short term as evidenced above, but it might come at the detriment of their productive asset base. For example, uninsured agricultural households losing a large portion of their harvest may be forced to sell livestock or other productive capital to be able to meet their needs in the short-run, but this may permanently lower income generation in the long-run (Heltberg, Oviedo and Talukdar, 2015). To examine if agricultural households in this context are drawing down on livestock in response to the 2012 flood to maintain consumption, I examine livestock impacts as the outcome variable. There are two dependent variables from the GHS-Panel that I use to measure livestock changes including: (a) total livestock value (in Naira) and (b) the value of livestock sales (in Naira). Additionally, I explore heterogenous impacts on consumption, livestock value, and sales for households below and above the poverty line to examine potential threshold effects (Carter *et al.*, 2007).

The main set of results on livestock outcomes are presented in Table 31. For livestock value (Column 1), flood-affected households experience a decrease in value by around 37,000 Naira, which represents 17% of the mean value but is not statistically significant. This drop could indicate livestock mortality due to the flood, as evidenced by Nzeribe, Nwokoye and Ezenekwe (2014). For the value of sales (Column 2), the coefficient in Column 4 is not statistically significant and suggests agricultural households impacted by the flood may not have sold productive livestock to maintain consumption.

Table 31. Impact of flood on livestock value and livestock sales for agricultural households.

VARIABLES	(1) Livestock value	(2) Value of sales
Flood	-37,258	-1,495
	(28,135)	(1,536)
Observations	768	768
Number of hhid	384	384
Controls	No	No
Household FE	Yes	Yes
Time FE	Yes	Yes
Cluster	HHID	HHID

Note: Livestock value and value of sales are represented in Naira. Results include household fixed effects and time fixed effects and are clustered at the household level. W95 indicates the dependant variable is top-end windsorized at the 95th percentile. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Not all households may be able to smooth consumption without depleting productive assets. In particular, households living in poverty or close to the poverty line may have difficulty maintaining consumption in the face of a shock (Baez *et al.*, 2017) or may be forced to sell livestock to stabilize consumption (Zimmerman and Carter, 2003). If households sell livestock in response to a shock, this may result in long-term poverty traps (Carter *et al.*, 2007). Here, I examine consumption and livestock changes for poor and non-poor households separately.

Table 32 and Table 33 investigate the heterogenous results by poverty status to examine the role of potential poverty thresholds (Carter *et al.*, 2007). For consumption (Table 32), both in terms of total and food-specific consumption, poor households affected by the 2012 flood experience a three-to-four times greater loss (comparing Columns 2 and 4 to Columns 1 and 3). For livestock impacts (Table 33), poor households similarly experience a much greater loss of livestock (Column 2). This may reflect the lack of resources for poor households to protect their livestock in the face of the flood. In terms of livestock sales, while the point estimate is negative for non-poor households (Column 3) and positive for poor households (Column 4), both coefficients are not statistically significant. As a result, there is no evidence of "distress sales" whereby households draw down on livestock to maintain consumption, with potential long-term impacts on the productive capital of households.

Table 32. Heterogenous impacts on consumption by poverty status

	(1) Nonpoor – Total consumption	(2) Poor – Total consumption	(3) Nonpoor – Food consumption	(4) Poor – Food consumption
Flood	-12,685*	-44,897***	-11,297**	-34,157**
	(6,623)	(6,192)	(4,929)	(4,887)
Obs	768	768	768	768
N of HH	384	384	384	384
Controls	No	No	No	No
HH FE	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes
Cluster	HHID	HHID	HHID	HHID
Outliers	W95	W95	W95	W95

Notes: Poverty threshold from the World Bank is used (49,994 Naira per capita). Results include household fixed effects and time fixed effects and are clustered at the household level. W95 indicates the dependant variable is top-end windsorized at the 95th percentile. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 33. Heterogenous impacts on livestock value and livestock sales by poverty status

	(1) Nonpoor – Livestock value	(2) Poor – Livestock value	(3) Nonpoor – Livestock sales	(4) Poor – Livestock sales
Flood	-28,875	-81,855**	-2,593	1,246
	(26,561)	(36,370)	(1,788)	(1,970)

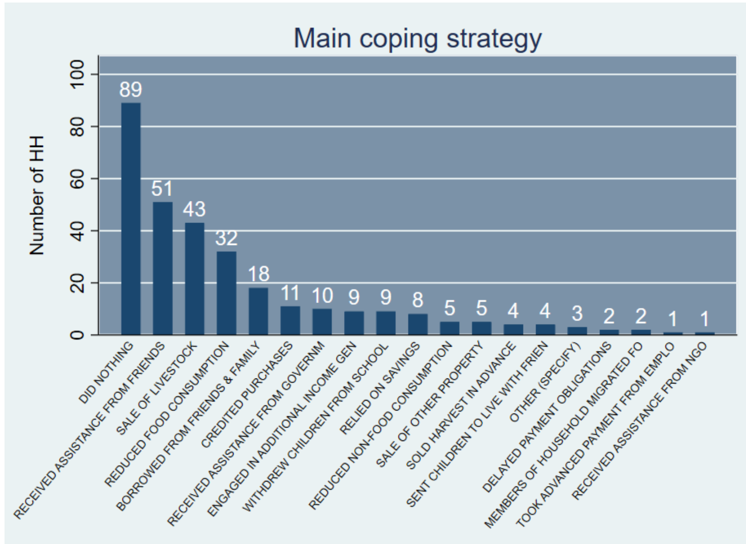
Obs	768	768	768	768
N of HH	384	384	384	384
Controls	No	No	No	No
HH FE	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes
Cluster	HHID	HHID	HHID	HHID
Outliers	W95	W95	W95	W95

Notes: Poverty threshold from the World Bank is used (49,994 Naira per capita). Results include household fixed effects and time fixed effects and are clustered at the household level. W95 indicates the dependant variable is top-end windsorized at the 95th percentile. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

4.4.3.2. Household coping strategies

In addition to the quantitative analysis using data from the GHS-Panel, the shock module also asks households what their main coping strategy was in response to the 2012 flood event. The results are presented in Figure 29. For the sample of all 300 affected households (244 rural and 56 urban), the most common coping strategy was to "Do nothing", with 89 households reporting this choice in the survey. This is followed by receiving assistance from friends (51), sale of livestock (43) and reduced food consumption (32). The 17% of households reporting livestock sales and 14% stating reduced food consumption provide qualitative estimates to some of the quantitative results found in this paper. One finding is that only 3% of households report receiving any assistance from the government or NGOs, indicating a lack of a formal safety net for households to manage risk.

Figure 29. Main coping strategy employed by households (both rural and urban) in response to the 2012 flood.



4.4.3.3. Responses in the flood-impacted village of Osomori

To complement the evidence gathered from the household survey, I conducted an interview with an individual whose family was directly affected by the 2012 flood. The respondent's family lived in the village of Osomori in Anambra state, residing on the banks of the Niger river in the southern part of the country. Osomori has very fertile soil and as a result its economy is largely agricultural, with many households employed in farming and fishing. It was one of the most severely impacted areas from the 2012 flood.

The interviewee first provided context on the relationship between people and water in Osomori. During the rainy season every year, the river swells to its largest capacity around the end of September, and starts to recede in the beginning of October. The water rises over the banks temporarily during this time period due to rainfall, and brings beneficial nutrients to the soil, but rarely ever goes into the villages where people live. In 2012, despite increased rainfall from July onwards, the water as usual rose over the banks but did not enter the village. However, in mid-September of 2012, government officials made the decision to release water from the overflowing Kainji Dam, which is located in the North West of Nigeria on the banks of the Niger river. As a result of this dam release, the water flowed into village communities, with half of the homes partially or fully destroyed, especially for households at lower elevation.

The intrusion of water into the community had severe impacts on agricultural livelihoods. Small-scale fisheries are a large source of income but were completely decimated as almost all of the fish escaped due to the flood. Due to the long duration of water intrusion and the velocity of the floods, the brick and cement used to keep the fish in farming ponds were weakened and completely broke in many cases. For example, the interviewee's brother had 40 ponds of fish, which all overflowed, and he was left with no income during that season. Agricultural lands were similarly inundated, with many farmers unable to harvest that season. The level of water intrusion was so severe that wild animals, including crocodiles, came out from the river and entered the community.

In terms of socio-economic impacts, schools were completely inundated and had to be closed for an extended duration, which impacted the human capital of children. With many households also losing their homes and having agricultural fields inundated, people temporarily fled to nearby camps where conditions can be challenging. In these camps, little governmental assistance is provided, food and shelter can be hard to come by, and employment opportunities are minimal. Income losses were so severe and prevalent, with many households relying on remittances alone to survive the season. A portion of the younger population decided to permanently relocate to the closest city, but conditions were tough in urban areas due to limited employment opportunities and the higher cost of rent.

In addition to the severe direct impacts of the flood in 2012, the interviewee also noted that recovery was slowed due to challenges for the 2013 growing season. In a usual year, farmers begin to plant for the next season as soon as the water recedes in October. However, due to the large-scale inundation, water levels did not recede until January, shortening 2013's growing season by 2-3 months. The main staple crop in Osomori is cassava, which is planted in October, takes about 8 months to mature, and is harvested before the area floods during the rainy season (in June-July). Despite the shorter growing season in 2013, farmers (whom the interviewee referred to as "stubborn") planted the same variety of cassava in early January. The farmers still had to harvest in June/July before the usual inundation arrives, and the cassava that was picked out of the ground was not mature. Farmer revenues suffered yet again, and many who had to take out loans after the 2012 production collapse went completely bankrupt. Due to supply shortages, food prices increased significantly for two years in a row and households had difficulty maintaining their food consumption. Two years of income loss, food unavailability, and forced migration led to social unrest due to the hardship.

Overall, the qualitative evidence obtained from the experiences in Osomori suggests that households were severely impacted by a human decision to release water from a dam, received little compensation or support from the government or other sources, and had a limited ability to respond and recover the following year. The interviewee suggested future interventions to improve recovery, especially in terms of farmer communication. Agricultural extension from the government can help farmers plant shorter duration crops (like rice, which can be planted 3x a year), potatoes, or vegetables if the following growing season is cut short due to inundation. These interventions may be paired with behavior change to adapt preferences in a post-disaster context (with the interviewee mentioning that people are used to eating cassava, which takes a long time to grow). Along these lines, better methods of storing cassava (e.g. fermenting and placing in rice bag covered with leaves) may be an option to mitigate the potential supply shortages in the years after a flood.

4.4.4. Robustness checks

The regression results presented above have not included any control variables. This section examines the robustness of the results on crop production, value, consumption, and livestock changes to including control variables in the analysis. As discussed in Section 4.3.2.1, several control variables that vary over time and may affect the relationship between floods and agricultural outcomes include the number of adults, bank account presence, landholdings, distance to market, and slope (McCarthy *et al.*, 2018). The main results for Table 21, Table 22, Table 23, Table 30, and Table 31 with controls is shown in the Appendix D2. The results with controls from Appendix D2 are almost identical to the results found without controls presented in the main body of the text and which provide evidence of the robustness of the main findings.

An additional robustness check that is run relates to the distance to market variable. In Table 20, the treatment and control groups show many similarities before the flood, but one difference is based on the distance to the nearest market, with the treatment group having to travel 12km further compared to the control group. A longer distance to market may be correlated with an under-investment in infrastructure generally in areas where the treatment group resides, which might increase flood exposure and impact agricultural and consumption outcomes. While the results do not change when controlling for distance to market, I also run a separate matching exercise to account for the possibility that distance to market is confounding the main results of interest.

The matching exercise matches households based on the local government authority (LGA) (equivalent to a sub-district) in which they reside. I restrict the sample to LGA areas that included flood inundation (at the 3km buffer) but where some but not all households within the LGA directly report being flooded. This removes LGAs where all households report being flooded, or no households report being flooded. From this matching exercise, 44 households remain in the treatment group and 138 in the control group from 27 LGAs across the country. The descriptive statistics are displayed below in Table 34. Here, the difference between the households in terms of distance to market (and age of the household head) disappears, as the two groups are statistically comparable before the flood across all characteristics.

Table 34. Household characteristics before the flood for treatment and control groups.

Characteristics	Treatment	Control	T-stat (diff)	P-value
No. of adults	3.64	3.65	0.05	0.96
HH head age	48.23	50.95	1.10	0.27
Bank account	0.09	0.13	0.70	0.49
Poverty status	0.48	0.43	-0.49	0.62
Landholdings	2098.96	2048.73	-0.12	0.91
Distance mkt	72.41	63.69	-1.29	0.20
Slope (%)	2.14	2.19	0.25	0.80
Elevation (m)	227.11	289.75	1.37	0.17
No. HH	44	138		

Note: Landholdings and the distance to market are represented in square kilometres.

The main results are then replicated in Appendix D3 and are broadly similar to the main results presented in the paper. However, due to much reduced sample size, the coefficient representing the impact of the flood on total agricultural production is no longer significant. Similarly, when exploring spillovers, households who live in a satellite flood zone and self-report being affected still lose crop value from the flood, but similarly the coefficient is not statistically significant. Nonetheless, the main results on crop value, the crop value impact on households in satellite flood areas but individually unaffected, consumption, and livestock remain unchanged.

4.5. Conclusions and policy implications

The 2012 floods in Nigeria were unprecedented, led to over 431 fatalities, displaced almost 1 million people, with economic damages estimated at $16.9 billion USD. However, this does not tell the full story on the impact on the livelihoods of the poorest, which is a focus of this paper. This paper examines the exposure, vulnerability, and ability to respond of households to an extreme flood event and makes three main contributions. First, I provide a more comprehensive accounting of flood impacts in sub-Saharan Africa by exploring local household outcomes across Nigeria for each component of the risk framework: exposure, vulnerability, and ability to respond. Second, methodologically, I combine satellite imagery and household self-reported data examine the causal impact of the flood. I also combine this quantitative analysis across the country with an interview with an individual impacted in a severely impacted village. Third, I examine spillovers and interactions with local market factors which uncover nuanced relationships between households and flood risk in a sub-Saharan African context.

By combining household survey data with spatial data before and after the flood, I find affected farmers lost around 20% of crop production and 50% of value on average. Additionally, impacts depend on the interaction between the individual and area-level measures of flood exposure. While farmers who live in a flooded area and directly experience the flood lose half their crop value, those in flooded areas but not directly affected experience increases in crop value by around 30%, which is related to food price increases. I also find that households living outside of historically flooded zones experience more severe impacts. In terms of household responses, I do not find evidence of consumption or livestock changes for the full sample, but poor affected households do experience drops in consumption and livestock value. This quantitative analysis across the country is complemented with a evidence from the village of Osomori, which finds that households face several constraints to their recovery and may not receive adequate support to foster adaptation.

It is important to consider several limitations of this paper when interpreting the results. First, on flood exposure, measurement error may be present both on the location of the household (whose coordinate is offset randomly by 5km) as well as for the flood extent produced by the satellites (which do not account for cloud cover). In terms of the self-reported data, the wording of the questionnaire does not allow the ability to capture flooding that did not "cause harvest failure", and the responses themselves may be prone to human bias. Second, on agricultural impacts, this study did not collect fine-resolution data on the differential timing of planting and harvesting across different regions of the country, which can determine overall agricultural impacts. Third, I do not observe data on farm-level practices to reduce the impacts of floods or other household-level adaptations to shocks such as the ability to store grain during a lean season. Forth, this paper only examines the short-term impacts of the 2012 flood, and not medium or long-run changes in household outcomes and recovery.

Future research can build on this paper to address the above limitations and better understand flood risk at the local level in Nigeria and across sub-Saharan Africa. First, collecting data on the cropping calendar and the exact timing of inundation at the very high resolution to include in a panel regression framework can help identify more causal effects of flood impacts (Guiteras, Jina and Mobarak, 2015; Baez *et al.*, 2017; Shiru *et al.*, 2019). Additionally, linking production and market data to separate crops that are locally traded from those that are internationally traded can help better understand market responses and heterogenous impacts for farmers impacted by flood shocks. Furthermore, the interviewed identified fish income as being particularly important, but may be overlooked in traditional agricultural analyses. Efforts to collect better data on aquaculture can better estimate total livelihood impacts of floods.

Second, this paper has shown that the same flood event can be measured very differently, with outcomes similarly varying. There are a plethora of options to measure floods, either through household surveys, flood risk models, or from space. Even within the realm of satellite data, 11 different (public and private) satellites currently exist to measure floods, each with its limitations and advantages (Tellman, 2023). When studying extreme events, it may be important to clearly define and state how exposure is classified, and to collect as many local details as possible to understand heterogenous effects and distributional impacts especially if certain thresholds are set which trigger insurance payments. Third, two more waves of the GHS-Panel exist for Nigeria (in 2015 and 2019), which can be used to explore longer-run outcomes of the 2012 flood, impacts of repeated flood events, and options for rural adaptation and labor reallocation especially for households who live outside of historically flooded areas (Blakeslee, Fishman and Srinivasan, 2020). Forth, the evidence collected from the interview conducted suggests that dam releases exacerbated the heavy precipitation event. Future research may want to model flood inundation from dam releases and precipitation separately, better understand the policy process through which decisions get made on water releases, and uncover whether communities downstream are informed.

There are also several policy implications that may be of consideration for decision-makers working on floods in Nigeria and across The Global South, especially as flood impacts are likely to worsen due to climate change. While flood risk may not have historically ranked high amongst the government's list of priorities, the devastating floods this past year in 2022 have spurred a renewed call to action to prioritize disaster risk management (Al Jazeera, 2022). Some of the institutional infrastructure is already in place: The Nigerian Meteorological Agency warned in March 2012 of severe floods during the rainy season, but little action was taken to prepare (NIMET, 2012). The same situation repeated itself in 2022 (Al Jazeera, 2022). An immediate policy change could be to better communicate early warning signals to reduce the loss of life and property damage in the aftermath of a flood. In terms of post-disaster support, there are number of options that countries at a comparable income level across The Global South have enacted including cash transfers, insurance schemes, and market interventions to release grain reserves and stabilize food prices (Hallegatte *et al.*, 2020). Future research can help inform which policy is best suited to the institutional setting. Finally, the case study of Osomori suggests that improved agricultural extension in the aftermath of a flood can help households plant shorter-duration crops to enable recovery.

Chapter 5: The impacts of fuel conversion on households: An assessment of Indonesia's Kerosene-to-LPG program in West Nusa Tenggara province

Abstract: Improving household access to cleaner cookstoves has long been a priority for governments across The Global South. One country which has been a leader is Indonesia, which in 2007 embarked on an LPG Conversion Program to convert 50 million households from using a kerosene stove to using liquefied petroleum gas (LPG), which is cleaner-burning and more fuel-efficient for cooking. This paper adds to the growing literature on the LPG Conversion Program by focusing on local impacts in the province of West Nusa Tenggara. Using a causal difference-in-difference framework with panel household survey data, I exploit the differential timing of implementation between 2 districts within this province to assess the short-term impacts on spending patterns and health outcomes. I conduct a policy review to show that while the two districts were scheduled to receive the program around the same time, difficulties in program implementation in other parts of the country delayed implementation in Sumbawa for four years. Under this empirical set-up, Lombok is considered the treated group and Sumbawa is a plausible counter-factual and serves as the control group, and the intention-to-treat of the program is uncovered. Program take-up is strong with over 60% of treated households report using the new LPG stove as their main cooking modality 4 years after the program, which resulted in substantial savings in fuel expenditures of around 13%, with higher benefits accruing to poorer households. While economically meaningful, savings represent less than 1% in terms of total expenditures, and I do not find evidence that households reallocated consumption towards pro-social investments. In terms of health, I do find some evidence that women report fewer illnesses related to indoor air pollution, but the results may be prone to measurement error. The findings of this paper suggest that it may be difficult to deliver multiple benefits beyond fuel savings to households in the short-term, and have implications for the design of similar fuel conversion programs currently being planned in a number of countries across The Global South.

5.1. Introduction

Improving household access to cleaner cookstoves has long been a priority for local and national governments, and was in 2015 enshrined in the Sustainable Development Goals of the United Nations (UN, 2016). Household use of improved cookstoves burning liquefied petroleum gas (LPG) is associated with lower indoor air pollution and improved health outcomes while also reducing fuel expenditures, compared to cookstoves using fuels such as biomass or kerosene (Gordon et al., 2014, Patel et al., 2015). Despite these benefits, households themselves are often reluctant to switch to cleaner cookstoves, due to their cost and unavailability, but also due to preferences for existing stoves (van der Kroon et al., 2013). This lack of autonomous adoption from households has opened up the role for governments to enact policies to incentivize conversion to cleaner cookstoves (Smith et al., 2014).

Indonesia is one such country that has intervened to incentivize households to use cleaner cookstoves. Starting in 2007, the government embarked on its *LPG Conversion Program* (also known as the Conversion Program), an initiative to convert 50 million households from using kerosene or biomass to cleaner-burning LPG for cooking, in the largest program of its kind. Under the initiative, each household receives a free LPG starter kit, including a gas stove, a 3-kg LPG canister, and all necessary connections (Permadi et al., 2017). The driving force behind the policy was to reduce the ever-growing government subsidy bill for kerosene, which totaled $3.8 billion in 2006, and promote LPG use (Toft et al., 2016). In addition, the program was expected to deliver several benefits to households, chief among them improved health outcomes and reduced fuel expenditures (Wiratmaja, 2016).

This paper builds on existing non-experimental (Budya and Arofat, 2011; Andadari, Mulder and Rietveld, 2014; Thoday *et al.*, 2018) and quasi-experimental (Imelda, 2018, 2020; Verma and Imelda, 2023) research exploring the impacts of the *Kerosene-to-LPG Conversion Program* by focusing on program experiences in West Nusa Tenggara province. Using a causal difference-in-difference estimation strategy, I exploit the differential timing in program implementation across two districts (Lombok and Sumbawa) within the province. I conduct a policy review to show that while the two districts were scheduled to receive the program around the same time, difficulties in program implementation in other parts of the country delayed implementation in Sumbawa for four years. Under this empirical set-up, Lombok is considered the treated group and Sumbawa is a plausible counter-factual and serves as the control group, and the intention-to-treat of the program is uncovered.

The first contribution of this paper is to examine socio-economic and health impacts of the program in one province of the country, with the results potentially being more internally valid than previous studies. There are a few reasons why this may be the case: first, I conduct an in-depth policy review of implementation within West Nusa Tengarra to provide a narrative for why program timing was plausibly random, with one district included in the program years before the other. Second, this paper examines impacts within one geographical setting where unobservable characteristics are more likely to be similar. Third, while previous work compared outcomes for early vs. late treated, this paper compares outcomes for treated vs. untreated, reducing the potential for anticipation to bias the results. In addition to internal validity, while previous work has examined impacts in early phases of the program, this paper focuses on implementation in later phases and expands the set of health outcomes studied.

Compared to many other studies in the literature on improved cookstoves, reported take-up is strong in this context. Four years after implementation, over 60% of treated households report using the new LPG stove as their main cooking modality. I find that sustained use of LPG through the Conversion Program led to substantial fuel savings of around 13% for treated households which is robust across samples and specifications, with benefits accruing progressively to poor households. While these savings are economically meaningful, they account for less than 1% of total expenditures, and I do not find evidence that households reallocate consumption towards pro-social spending like education, In terms of health outcomes, I do find evidence that women included in the program do report fewer illnesses related to indoor air pollution, but this result is small in magnitude, may be prone to measurement error, and does not measure particulate exposure at the household level. The findings of this paper suggest that it may be difficult to deliver multiple benefits beyond fuel savings to households in the short-term and have implications for the design of similar fuel conversion programs currently being planned across The Global South.

The remainder of this paper proceeds as follows. Section 5.2. provides background on the Conversion Program and conducts a literature review of related studies in Indonesia and across The Global South. Section 5.3. outlines the research design. Section 5.4 presents the main results and Section 5.5. assesses robustness and limitations of the study. Section 5.6. concludes and provides implications for policy.

5.2. Program background and related literature

Indonesia began to implement the *Kerosene-to-LPG Conversion Program* starting in 2007. The major motivation was fiscal: kerosene had been subsidized by the government since the 1970s, but along with the rising population and oil price, the subsidy bill became a large burden (Toft et al., 2016). In 2006, the annual subsidy bill for kerosene reached $3.8 billion USD (PT Pertamina, 2013). For the end user, LPG is almost three times more efficient fuel compared to kerosene (Budya and Arofat, 2011), and a large-scale shift to convert 50 million households would likely produce substantial savings. Estimates from the government suggest that from 2007-2016, the *Kerosene-to-LPG Conversion Program* saved $15.6 billion USD (Wiratmaja, 2016).

Under the initiative, each household receives a free LPG starter kit, including a gas stove, a 3-kg LPG canister, and all necessary connections (Permadi et al., 2017). The program was implemented in four phases. Early policy documents released in 2013 show that Western Java and Sumatra were treated in Phase 1 (2007-2008), and other parts of the country in subsequent phases: Phase 2 (2009), Phase 3 (2010-2011), and Phase 4 (2012-2013) (PT Pertamina, 2013). The areas which were selected first for the program were those which had high kerosene consumption and adequate LPG infrastructure and supplies (Andadari et al., 2014).

Due to the large-scale nature of the program, it has attracted the attention of researchers and government evaluators alike. In general, the program has been viewed as a success story, delivering multiple socio-economic and health benefits to households in addition to fiscal savings (Pertamina, 2012; Toft, Beaton and Lontoh, 2016; Imelda, 2020). Government communication and consumer surveys highlight the potential benefits delivered to households in terms of fuel savings, reallocation of expenditures towards pro-social consumption like education, and improved health due to reduced indoor air pollution (Pertamina, 2012; Toft, Beaton and Lontoh, 2016; Wiratmaja, 2016). Below, I review the evidence on (1) take-up rates, (2) fuel savings, (3) reallocation of expenditures, and (4) health benefits of Indonesia's program while including studies from the literature across The Global South to provide context.

On the first question, follow-up government surveys and evidence from the literature finds strong take up rates. One year after implementation, official surveys indicate take-up rates of the LPG cookstove of around 70-80% by households (PT Pertamina, 2013). Academic studies find similar rates of adoption: Latifah (2010) examines take-up rates of low-income households in the city of

Bogor. Before the intervention, 45% of households reported using kerosene as their main cooking fuel, and 55% reported using firewood. After the program, the composition changed remarkably, with 82% reporting LPG as the main cooking fuel, and only 10% and 8% using kerosene and firewood, respectively. Budya & Arofat (2011) present the results from a consumer satisfaction survey undertaken among 550 households, and report 86% of respondents preferred LPG to kerosene after the program. Evidence suggest these take-up rates continued to remain strong: across the country, total kerosene consumption in dropped dramatically by 92% and LPG household consumption increased five-fold from 2006 to 2015 (Thoday *et al.*, 2018). In terms of shares, the proportion of households using LPG jumped from 9% to 46% while the proportion of households using kerosene dropped from 42% to 12% (Imelda, 2018).

The high-rates of adoption of more efficient cookstoves in Indonesia are in contrast to existing evidence from individual case studies in the literature. For example, a study in Orissa, India, finds that households generally do not use improved stoves, with take-up and usage declining markedly even one or two years after receiving it (Hanna et al., 2016). Households in this context had a preference for traditional stoves and failed to make the maintenance investments in the improved stove to keep them fully operational. Similar evidence of low take-up has been found in Ghana (Burwen and Levine, 2012) and Bangladesh (Miller and Mobarak, 2013). The context of Indonesia is slightly different, where the nationwide intervention included changing both the stove and the type of fuel used. Additionally, the policy roll-out to encourage LPG was paired with a subsequent removal of kerosene from areas that were included in the program (Pertamina, 2012), with the lack of availability of alternative fuels potentially explaining the high adoption rates witnessed in Indonesia (Imelda, 2018). The closest counterpart to Indonesia in The Global South is Ecuador, which enacted a subsidy for LPG starting in the 1970s (Gould *et al.*, 2018). However, in Ecuador's case, kerosene was unsubsidized and at low levels of use, with LPG primarily substituting for wood use (Hakam *et al.*, 2022). The program, while costly, was highly popular politically and experienced high take-up rates, with adoption increasing from 5% before 1980 to almost 90% as of 2014 (Gould *et al.*, 2018). In Ghana, a government program to incentivize rural households to use LPG increased consumption of the fuel by 24-33%, with higher adoption related to shorter distance to refill markets (Adjei-Mantey, Takeuchi and Quartey, 2021).

Related to take-up, if households transition from using kerosene to LPG for most of their cooking, they may experience savings in fuel expenditure due to the higher efficiency of LPG and reduced cost (Toft, Beaton and Lontoh, 2016). Early evaluations do suggest that households included in the program experienced lower fuel expenditures by up to 30% (Budya and Arofat, 2011; Pertamina, 2012; Andadari, Mulder and Rietveld, 2014). However, these studies are either based on government follow-up surveys (where households might have an incentive to report favorable outcomes) or evaluations that only focused on households who were treated and did not examine outcomes for unaffected households.

More recent quasi-experimental evidence from a working paper exploits the differential timing of the program across Indonesia and finds significant savings in fuel expenditures (Imelda, 2018). Comparing outcomes for early treated and late treated households using panel data across the country finds that households in targeted districts reduced kerosene consumption by up to 100% and experienced savings in fuel expenses by up to 40%. Outside of the Indonesian context, prior studies which documented low-take up in India, Bangladesh, and Ghana unsurprisingly do not find any evidence of fuel savings (Burwen and Levine, 2012; Miller and Mobarak, 2013; Hanna, Duflo and Greenstone, 2016).

If households save on fuel expenditures, it may impact overall spending patterns. For instance, households may be able to reallocate their consumption towards human capital investments in health or education. While there is less evidence on how households reallocate savings in response to a "positive shock", the extensive literature on how households respond to a negative shock can be informative. For example, in responses to natural disasters, studies often find that consumption on education reduces as a result of the shock (Dercon, Hoddinott and Woldehanna, 2005; Akter and Basher, 2014; Hallegatte, Bangalore, *et al.*, 2016). As a result, we might expect a similar effect but for a positive "shock", with expenditures for education increasing. In the quasi-experimental study examining the impacts of Indonesia's Kerosene-to-LPG Program on household consumption, Imelda (2018) finds that fuel savings accumulated to 2% of overall spending, which did not appear to be reallocated towards human capital investments.

The other major potential benefit of shifting towards LPG for households commonly cited in government documents are improved health outcomes related to less indoor air pollution (Pertamina, 2012). There is agreement in the literature that reduced indoor air pollution leads to better health outcomes (Epstein *et al.*, 2013; Kurt, Zhang and Pinkerton, 2016; World Health Organization, 2022) and the World Health Organization has recently classified kerosene as a dirty and polluting fuel (Bryden, Bruce and Peck, 2015). Thus, sustained adoption of LPG cookstoves can reduce indoor air pollution by decreasing the amount of particulate matter and other emissions if it replaces kerosene (Thoday *et al.*, 2018). Additionally, on the intensive margin, cooking with LPG is 40% faster compared to kerosene, which might further reduce the overall levels of indoor air pollution (Pertamina, 2012).

The literature across The Global South exploring the health impacts of clean cooking intervention is extensive, with many studies using randomized control trials to identify impacts. In each context, each intervention is slightly different. Many studies distribute new technology that uses the same fuel as before (e.g. firewood) but with improved ventilation, while the intervention in other countries (like in Indonesia) changes both the fuel and stove being used. The focus of these papers has been on the health impacts for women and children, as women are typically responsible for cooking activities across The Global South and thus experience more indoor air pollution (Pitt, Rosenzweig and Hassan, 2005; Hanna, Duflo and Greenstone, 2016; Toman and Bluffstone, 2017). While children may have less exposure to indoor air pollution due to schooling, they are more vulnerable to health impacts for a given level of exposure (Smith et al., 2014).

The evidence from the literature provides mixed results in terms of health impacts. In Guatemala, an intervention to provide households with an improved cookstove with a chimney did lead to reduced blood pressure readings for women (McCracken *et al.*, 2007). In Nigeria, Alexander et al., (2017) finds that women receiving a cookstove with ethanol fuel (compared to biomass) similarly reduced diastolic blood pressure during pregnancy. In terms of impacts on children, the same intervention in Guatemala produced mixed results: overall pneumonia risk for children under 18 months did not change, but diagnoses of severe pneumonia dropped by a third (Smith *et al.*, 2011). Relatedly, an intervention in rural Malawi to provide households with cleaner-burning biomass stoves found no evidence of a reduction in pneumonia risk for young children (Mortimer *et al.*, 2017). In Orissa, India, a large-scale randomized trial found no changes across health outcomes of distributing clean cooking interventions, as households did not consistently use or maintain the cleaner stove (Hanna, Duflo and Greenstone, 2016).

More recently, several studies have examined the impact of switching from a dirty fuel (biomass or kerosene) to LPG. Both in Nepal and Ghana, LPG interventions reduced indoor air pollution significantly compared to traditional cooking methods (Katz *et al.*, 2020; Chillrud *et al.*, 2021) but did not lead to improved birthweight for infants or reduced severe pneumonia risk (Asante *et al.*, 2019; Katz *et al.*, 2020; Jack *et al.*, 2021). In both cases, while the transition to LPG did reduce indoor air pollution significantly, post-intervention exposures still exceeded health-relevant targets (Chillrud *et al.*, 2021). The lack of consistent health impacts found in these studies may not be surprising, as health is typically not the main focus of policy design to promote LPG, which is motivated more by economic, environmental, or gender concerns (Quinn *et al.*, 2018). However, some papers do find significant health benefits from reducing dirty fuel use and increasing LPG use. Adjei-Mantey & Takeuchi (2021), also in Ghana, finds that LPG interventions reduces in-utero exposure and increases child height after birth. In Indonesia, Silwal & McKay (2015) find that women and children in households that cook with firewood have a 9.4% lower lung capacity compared to households that cook with cleaner fuels (however, in this study, both LPG and kerosene are considered as "clean"). Barron & Torero (2017) study the removal of kerosene for lighting in El Salvador and finds a 66% drop in fine particulate matter and a reduction of 8-14 percentage points in respiratory infections for children under 6.

While the evidence from the literature suggests that transitioning to LPG is likely to reduce indoor air pollution, it may not lead to health improvements for three reasons. First, households may not maintain the new stove (Hanna, Duflo and Greenstone, 2016) or live close to an LPG refill station (Adjei-Mantey, Takeuchi and Quartey, 2021), reducing use over time. Second, households may use multiple cooking modalities including LPG, kerosene, and biomass (a phenomenon known as "fuel stacking"), due to preferences for traditional fuels associated with familiarity and taste (Toman and Bluffstone, 2017; Krishnapriya *et al.*, 2021). Third, households in developing countries tend to live in polluted environments, and it is unclear as to whether improvements from using an LPG stove are high enough to translate to health improvements (Jack *et al.*, 2021).

Specifically for the *Kerosene-to-LPG Conversion Program* in Indonesia, early evidence suggests the program led to a reduction in extreme energy-poverty among rural households, and may have improved health outcomes due to reduced air pollution (Andadari, Mulder and Rietveld, 2014). However, this study did not examine outcomes for households untreated by the program. Recent published studies employing quasi-experimental methods finds that the program resulted in a multitude of health benefits for women and children (Imelda, 2020; Verma and Imelda, 2023). Comparing early-treated to later-treated districts, women in early-treated districts experienced a 4% increase in lung capacity compared to women in later-treated districts, with no change for men (Verma and Imelda, 2023). This paper also explores self-reported health outcomes of cough, stroke, diabetes, and overall health. While the authors find a lower probability of self-reported cough symptoms, stroke, and diabetes, and a higher probability of reporting good health, most of the coefficients are not statistically significant (Verma and Imelda, 2023). Using a similar methodology, Imelda (2020) examines program impacts on children and finds a significant decline in infant mortality by around 16-34% and reduced prevalence of low birth weight by 8-25%. Both papers examine potential channels of impacts, and provide evidence that the mechanism through which health impacts materialize is through reduced indoor air pollution (Imelda, 2020; Verma and Imelda, 2023).

This paper builds on existing non-experimental (Budya and Arofat, 2011; Andadari, Mulder and Rietveld, 2014; Thoday *et al.*, 2018) and quasi-experimental (Imelda, 2018, 2020; Verma and Imelda, 2023) research exploring the impacts of the *Kerosene-to-LPG Conversion Program* by focusing on program experiences in West Nusa Tenggara province. The first contribution of this paper is to examine socio-economic and health impacts of the program in one province of the country, with the results potentially being more internally valid than previous studies. There are a few reasons why this may be the case: first, I conduct an in-depth policy review of implementation within West Nusa Tengarra to provide a narrative for why program timing was plausibly random, with one district included in the program years before the other. Second, this paper examines impacts within one geographical setting where unobservable characteristics are more likely to be similar. Third, while previous work compared outcomes for early vs. late treated, this paper compares outcomes for treated vs. untreated, reducing the potential for anticipation to bias the results. In addition to internal validity, while previous work has examined impacts in early phases of the program, this paper focuses on implementation in later phases and expands the set of health outcomes studied.

5.3. Research design

5.3.1. Research questions

Related to the four household-level outcomes of fuel conversion programs reviewed in the literature, I examine how the Conversion Program impacted (1) LPG use, (2) fuel and education expenditures and (3) health outcomes within West Nusa Tenggara province, comparing outcomes for households in Lombok to households in Sumbawa. Specifically, the questions explored are:

1. Did households in Lombok report changes in fuel use compared to households in Sumbawa?

2. Did households in Lombok experience savings in fuel expenditures compared to households in Sumbawa? Were potential savings reallocated towards pro-social investments in education and/or health?
3. Did women and children in households in Lombok experience any change in self-reported health outcomes or biometric indicators of lung health compared to women and children in Sumbawa?

5.3.2. Data

Data on policy implementation of the Conversion Program comes from government documents (Pertamina, 2012; Wiratmaja, 2016) and reports from the think-tank Global Subsidies Initiative (Global Subsidies Initiative, 2014, 2015). The main data source used to assess household adoption of the LPG cookstoves and track socio-economic outcomes over time is the Indonesia Family Life Survey (IFLS), conducted by the Rand Corporation (Strauss and Witoelar, 2022). These are detailed household surveys that collect information on multiple indicators of economic well-being, health outcomes, and demographics. Importantly, they are panel surveys and track the same households over time. In this paper, I use data from four waves of the survey, which took place in 1997 (Wave 2), 2000 (Wave 3), 2007 (Wave 4) and 2014 (Wave 5)[54]. Within West Nusa Tenggara, the Conversion Program was implemented in 2010 in the treated district of Lombok, which provides pre-treatment data at three points in time (1997, 2000, and 2007), and post-treatment data at one point in time (2014). Importantly, due to the delay in program implementation, Sumbawa did not receive the program until after the Wave 5 data which was collected in 2014.

The IFLS covers 13 of the 26 provinces in Indonesia, which represents 83% of the country's population. Significant resources are spent to track the same households over time, regardless of migration between waves. Specifically in West Nusa Tenggara, 440 households were surveyed in Wave 2 in 1997. Of these 440 households, 402 remained in the same district as of Wave 5 in 2014, representing an attrition rate of less than 9%. This panel of 402 households constitutes the main sample of interest, with 239 in the treated district of Lombok and 163 in the untreated district of

[54] The survey was started in 1993, but the wave implemented in that year did not include all the questions necessary for the analysis, so I start with Wave 2 from 1997.

Sumbawa. In addition to this panel, more households were added to the survey in 2007 and 2014. To increase statistical power, I also test alternative specifications using this larger sample.

The data includes several variables to measure cooking patterns, overall and fuel-specific expenditures, and health outcomes. Cooking patterns are measured using the question "What is the main fire/stove used for cooking", and respondents can either answer[55]: (1) Electricity, (2) Firewood, (3) Kerosene, or (4) LPG. Importantly, this measure only asks about the main stove used for cooking, and does not capture "fuel-stacking", whereby multiple cooking modalities are used by the household and which may be common in this context (Krishnapriya *et al.*, 2021).

The IFLS also includes detailed expenditure data on how households spend their money with the main outcome variables representing total (and per capita) monthly expenditures on fuel, education, health, food, clothing, and housing. The rationale for including these expenditure categories is to examine if households reallocate potential fuel savings into other areas of their portfolio, for instance investments in human capital like education.

Health outcomes from the IFLS are measured both from self-reported data and from biometric assessments. In terms of self-reported data, I am interested in health conditions related to indoor air pollution, which include difficulty breathing, cough, headache, and general health (Naz, Page and Agho, 2017). In the survey, data is collected for each member of the household (the head of household provides the responses for children under 15) on whether each member exhibited symptoms of "any illness", "headache", "cough" or "difficulty breathing" in the 4 weeks preceding the survey. I am interested in examining data separately for men, women, and children as in Indonesia, women are generally the members of the household most involved in cooking[56] and have higher baseline exposure (Thoday *et al.*, 2018) and children are more vulnerable for any given exposure (Mortimer *et al.*, 2017). Using these responses, the outcome variables of interest present the share of women, children, and men per household who report experiencing these symptoms. However, one limitation identified in the literature is that self-reported health measurements may be prone to measurement error. Household responses are subjective and some respondents are

---

[55] Households can also respond by reporting "Charcoal" "Do not cook", or "Other", but these responses contribute to less than 1% of the answers and are dropped from the main analysis.
[56] I confirm in the data that women take the primary responsibility for cooking, with 98% of households reporting that women are the primary cooks.

systematically more likely than others to report experiencing symptoms of illness (Bertrand and Mullainathan, 2001).

To account for this potential bias, I also introduce another outcome variable to measure health status that is based on observable data. As part of the survey, trained health workers conduct tests on members of the household, collecting biometric data on variables such as height, blood pressure, and lung capacity. I select the variable representing respiratory conditions – lung capacity – which is often used in the literature[57] (Pellegrino *et al.*, 2005; Verma and Imelda, 2023). Lung capacity is measured on a scale from 0-800 ml/kg, with higher scores representing better respiratory health. In the analysis, I present results for the raw lung capacity scores as well as an age-corrected score to account for the fact that younger people have generally higher scores[58]. I also collect data on electricity access, share of households with a male head, schooling of household head, and ownership of savings account, to examine potential heterogenous treatment effects and a range of other demographic and socio-economic variables to examine balance between samples.

5.3.3. Identification strategy

The empirical challenge in this context is to isolate the impact of the Conversion Program. There are often reasons to believe that treated and untreated units differ in unobservable characteristics associated with potential outcomes, even after controlling for observable characteristics (Angrist and Pischke, 2008). In these cases, the treated and untreated units may not be directly comparable. For instance, particular districts may have experienced the benefits of LPG before the program and lobbied the government to receive the treatment first. In this case, pre-existing levels of LPG use (perhaps driven by higher incomes) could be driving the relationship of interest, and ordinary least squares estimators might provide a biased (and overstated) estimate of the program's impact. Reverse causality may also occur if districts with high levels of respiratory illnesses or high expenditures on kerosene were prioritized by the program. In this case too, simple OLS estimates would result in biased estimates.

[57] The test is conducted by asking participants to exhale into a machine, and calculates the pressure produced from each participant.

[58] I calculate this metric by averaging lung capacity per age group (separately for women, men, and children), represented by 10 age quantiles. Then, the difference between each participant and the average score is calculated to estimate the age-corrected lung capacity score.

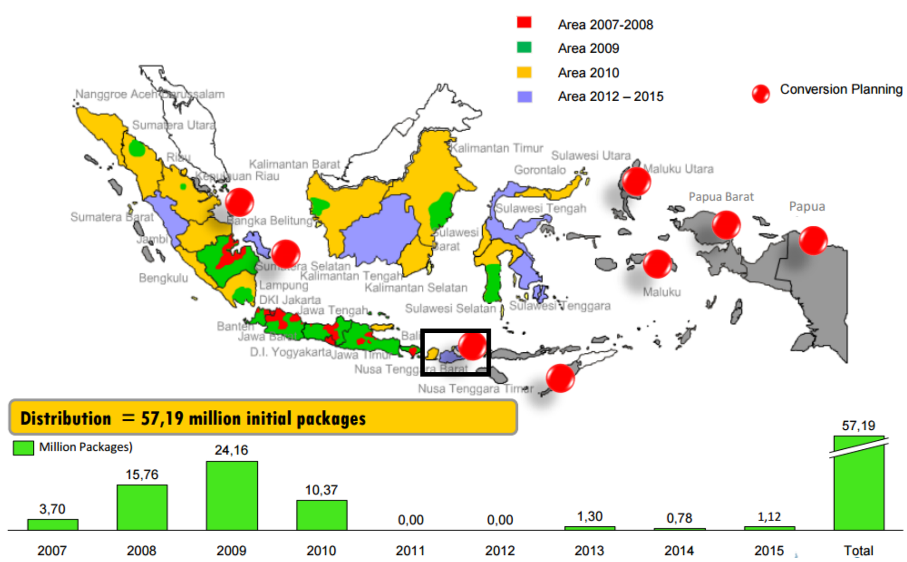
This paper builds on recent work employing quasi-experimental techniques (Imelda, 2018, 2020; Verma and Imelda, 2023) to isolate the impact of the Conversion Program by using the differential timing of implementation to compare outcomes for early-treated and late-treated households (Figure 30). To complement these studies which focus on the entire country, I examine program impacts in one province (West Nusa Tenggara) with plausibly exogenous assignment to treatment across districts with similar characteristics.

The Conversion Program was implemented starting in 2007 in four phases, and implementation was not randomized. Districts with the highest kerosene consumption and readiness for LPG were selected by the Ministry of Energy and Mineral Resources to be included in the program first during Phase 1 (2007-2008). Reviews of policy implementation suggest that while districts in Phase 1 were clearly different from other parts of the country, the difference between districts in Phase 2 (2009) and 3 (2010-11), and especially in Phase 3 and 4 (planned 2012-13), is less clear cut (Pertamina, 2012; Toft, Beaton and Lontoh, 2016; Wiratmaja, 2016). Furthermore, the roll-out of the program was highly spatially correlated indicating areas near to each other would be treated at the same time. Given Indonesia's island geography, each province experiences very different energy conditions and socio-economic circumstances (Rentschler and Kornejew, 2017).

This paper examines program impacts within a particular province – West Nusa Tenggara – where districts were treated at different times. The research design and identification strategy of this paper relies on the delay in reaching all districts of West Nusa Tengarra during the later phases of the program (Figure 31). Lombok – an island district in the western part of the province, was included in Phase 3 in 2010-11. Sumbawa – another island district located 5 miles east of Lombok – was originally planned to receive the program in 2012, but only received the program after 2014 (Pertamina, 2012; Wiratmaja, 2016). This unexpected break in implementation was due to implementation challenges in other areas of the country. Given the large-scale nature of the program, and being the first of its kind globally, the initiative hit roadblocks along the way. A policy review indicates that the program was completely stopped in 2011 and 2012 and Phase 4 was delayed until 2013-2015 due to challenges providing adequate LPG supplies for in "problem provinces" during Phases 2 and 3 (Global Subsidies Initiative, 2014, 2015; Wiratmaja, 2016). Instead of extending the conversion program as originally planned to other areas, the program implementers decided to focus on treated areas to ensure the LPG supplies were available (Global Subsidies Initiative, 2015). While the policy review did not identify the "problem provinces" which led to the stall in implementation, it is unlikely to be West Nusa Tenggara, since other provinces treated by the program in Phases 2 and 3 were much larger in size (Andadari, Mulder and Rietveld, 2014).

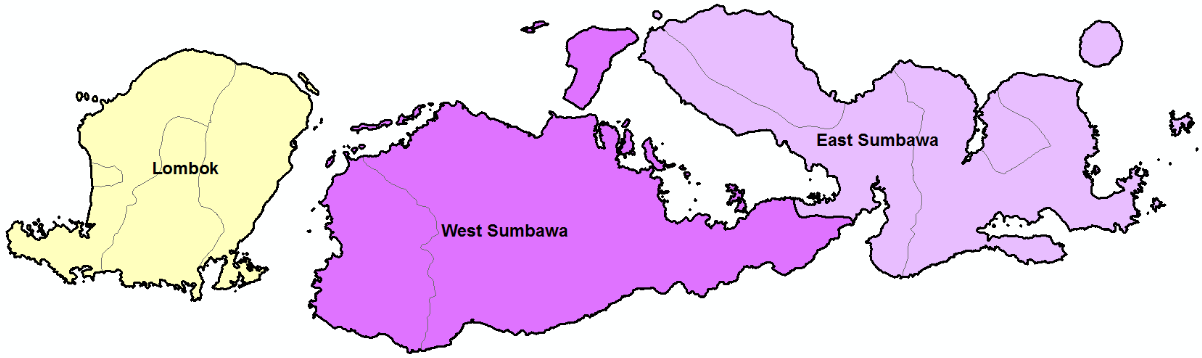
In summary, this paper compares the differences in socio-economic outcomes for households in Lombok district compared to Sumbawa district. While Lombok was selected in Phase 3 of the program and Sumbawa for Phase 4, Sumbawa can be seen as a counterfactual for two reasons: first, areas selected in later phases of the program have fewer differences than in earlier phases. Additionally, given Indonesia's island geography, these two districts are located within the same geographical province. Due to these similarities, combined with the unexpected delay in implementation in Sumbawa due to challenges in other areas of the country, this paper is able to identify the causal impact of the program on households in West Nusa Tengarra, by comparing outcomes in Lombok (treated) to those in Sumbawa (untreated).

Figure 30. Roll-out and progress of the Conversion Program in Indonesia as of 2015.



Notes: Based on (Wiratmaja, 2016). Highlighted black box in figure indicates study site, with close up in Figure 31.

Figure 31. Close up of study site within West Nusa Tengarra.



Notes: Based on (Wiratmaja, 2016). In Lombok, the program was implemented in 2010. In West and East Sumbawa, the program was implemented after 2014.

5.3.4. Empirical specification

This paper estimates the causal impact of the Conversion Program on household socio-economic and health outcomes through a difference-in-difference (DID) estimation strategy, which has been used in the quasi-experimental literature examining the Conversion Program (Imelda, 2020; Verma and Imelda, 2023). Such an empirical strategy has been used more broadly to evaluate the impacts of government programs in developing countries, for instance Mexico's conditional cash transfer program (Skoufias *et al.*, 2001), land titling in Argentina (Galiani and Schargrodsky, 2010), and environmental projects for reforestation in China (Groom *et al.*, 2010). The empirical specification is as follows:

Equation 5. Differences-in-difference strategy to estimate the impact of the conversion program on household outcomes

$$Y_{it} = \mu + \gamma_i + \delta t_t + \beta D_{it} + \varepsilon_{it}$$

Y_{it} is the socio-economic outcome of interest (LPG stove use, fuel and education expenditures, and health outcomes) for household i in survey wave t, γ_i represents household-fixed effects, δt_t time-fixed effects, and D_{it} indicates the treatment. The treatment variable of D_{it} uncovers the *Intention-to-Treat (ITT).* The *ITT* represents the impact of receiving the intervention, comparing households in Lombok (who were included in the program) to a plausible counterfactual of households in Sumbawa (whom were not included in the program). Additionally, Table 35 provides information on the terms used in this paper, their definitions, and the application in this research context.

Table 35. Definition of terms and application in this research context.

Term	Definition	Application in this context

Treatment	Implementation of conversion program	The program was implemented in Lombok and not implemented in Sumbawa (due to challenges in implementation elsewhere in the country).
Treated group	Households receiving the program	Households in Lombok (n=239)
Control group	Households not receiving the program	Households in Sumbawa (n=163)
Intention-to-treat	Estimator uncovered by the difference-in-difference analysis	Represents the impact of being included in the program (e.g. for households in Lombok)

5.3.5. Descriptive statistics

The IFLS data used in this paper contains three waves of the survey before the Conversion Program (Wave 2 in 1997, Wave 3 in 2000, and Wave 4 in 2007) and one wave after the program was implemented in Lombok (Wave 5 in 2014). This section explores trends in the data when it comes to the main fuel used, expenditure and health outcomes, and trends over time for the 239 households in Lombok and 163 in Sumbawa. Summary statistics for all variables is presented in Table 36, with explores trends over time in treated Lombok and untreated Sumbawa. The data is also presented graphically in Figure 32, Figure 33, Figure 34, Figure 35, and Figure 36.

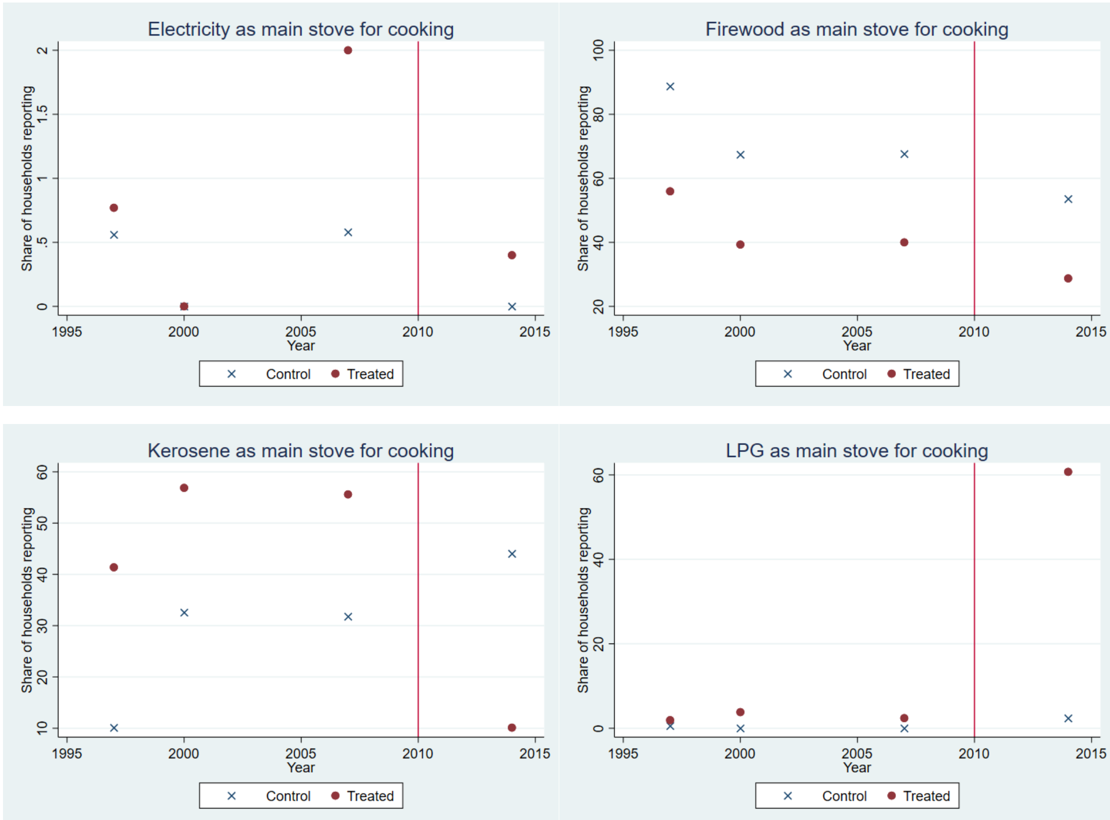
Table 36. Descriptive statistics for outcome variables in Lombok and Sumbawa across the four survey waves (1997, 2000, 2007, and 2014).

Variables	1997		2000		2007		2014	
	Lombok	Sumbawa	Lombok	Sumbawa	Lombok	Sumbawa	Lombok	Sumbawa
LPG use	1.92	0.56	3.82	0	2.4	0	60.73	2.38
Fuel exp	13,264	8,504	10,706	5,456	8,422	6,008	9,685	11,197
Educ exp	25,802	20,670	20,881	18,069	30,999	41,414	49,340	77,416
W-any	90%	75%	84%	76%	66%	53%	88%	82%
W-breath	8%	7%	5%	5%	1%	1%	2%	4%
W-cough	33%	25%	18%	18%	10%	6%	22%	21%
W-head	60%	43%	48%	37%	28%	23%	39%	34%
W-lung cap	-62.4	-50.8	4.0	-6.2	23.6	28.9	1.6	9.4
C-any	90%	82%	77%	68%	69%	61%	92%	87%
C-breath	3%	3%	1%	2%	0%	0%	2%	5%
C-cough	29%	20%	18%	17%	13%	5%	36%	32%
C-head	28%	20%	11%	11%	5%	6%	32%	28%
C-lung cap	-62.4	-41.6	-7.8	-12.8	13.3	28.7	0.0	-1.4
N of hh	239	163	239	163	239	163	239	163

Notes: LPG use indicates the share of households reporting LPG as their main stove of use. Fuel expenditures are total monthly fuel expenditures in real terms. Education expenditures are total monthly education expenditures in real terms. “W” stands for women and “C” stands for children. Any shows the share of women/children reporting any illness, Breath stands for breathing, and Head indicates headache. Lung cap measures the lung capacity, is age-corrected, and measured in mg/L. Lombok is the treated district and Sumbawa is the untreated district.

Figure 32 reports the trends in the main stove used for cooking and clearly shows the impacts of the program in the treated district of Lombok. Each of the four panels indicates the share of households using electricity, firewood, kerosene, and LPG. Electricity accounts for a very small percentage of responses (<2%) across time in both districts. Firewood accounts for a much larger share of responses, around 60% and 80% in Lombok and Sumbawa, respectively, starting in 1997. This trend declines over time, reflecting a general shift away from firewood to other sources. The panels related to kerosene and LPG are provided in the bottom-left and bottom-right. For kerosene, this stove is more important in Lombok (starting off around 40% in 1997) than Sumbawa (around 10% in 1997), but the two trends show identical patterns until 2007. However, the next data point from the IFLS is in 2014, when Lombok has been treated by the program (in 2010), and Sumbawa has not yet been treated. In 2014, the patterns diverge: kerosene decreases substantially from 60% in 2007 to 10% in 2014 in Lombok, but continues an upward trend in Sumbawa. LPG was used only minimally in both Lombok and Sumbawa from 1997 to 2007, and does not account for more than 2% in either district. However, the 2014 data reflects a different trend. Consistent with the decrease in kerosene in Lombok due to the Conversion Program, LPG increases from around 2% in 2007 to more than 60% in 2014. For Sumbawa, which was not treated by the time of the final survey round, the increase is only marginal, from 2% in 2007 to 5% in 2014.

Figure 32. Trends in main stove used for cooking across survey years.

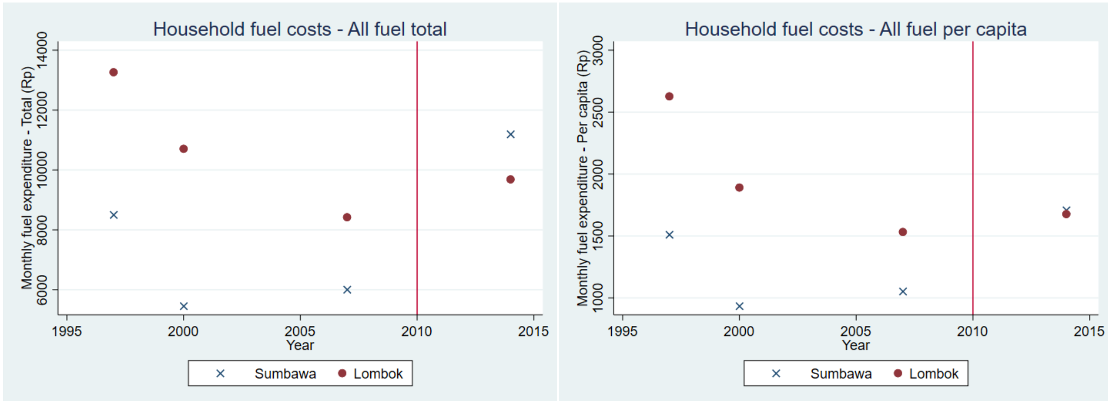


Notes: Based on IFLS. Treated indicates Lombok and control indicates Sumbawa. Red line indicates time of treatment in Lombok in 2010.

Data on inflation-adjusted fuel expenditures over time is presented in Figure 33. In 1997, households in Lombok spent 13,000 INR on fuel while households in Sumbawa spent around 9,000 INR on fuel, accounting for 2-3% of total expenditures in both districts. Accounting for inflation, and consistent with subsidized fuel provided by the government, fuel expenditures decreased in 2000 and 2007. The pre-treatment trends are very similar across the two districts but appear to increase more rapidly in Sumbawa (control district) after the treatment. Trends are similar if fuel expenditures are calculated on a per-capita basis, although the difference post-treatment between the two districts is slightly attenuated.

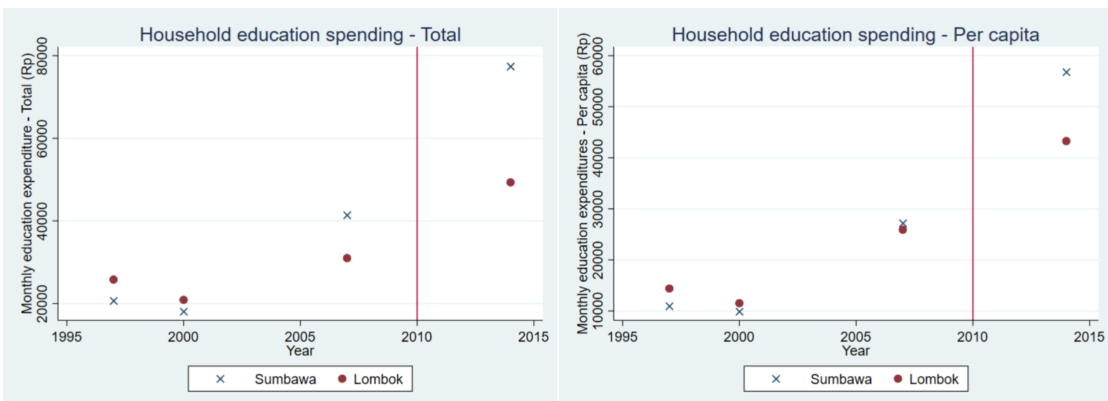
Trends for inflation-adjusted education expenditures are presented in Figure 34. In both districts, households on average spend around 1% of their total expenditures on schooling. Trends for education spending across survey waves also similar across Lombok and Sumbawa pre-treatment both on a total and per capita basis, although households in Sumbawa appear to spend more after the treatment. The same trends hold when examining expenditure shares on education, as well as other consumption categories including food, fuel, and medical expenses.

Figure 33. Household fuel expenditures across the survey waves, in total and per capita expenditures.



Notes: Treated indicates Lombok and control indicates Sumbawa. Data from 1997, 2000, and 2007 are pre-treatment. Data in 2014 is post-treatment. Red line indicates time of treatment in Lombok in 2010.

Figure 34. Household education expenditures across the survey waves, in total and per capita expenditures.



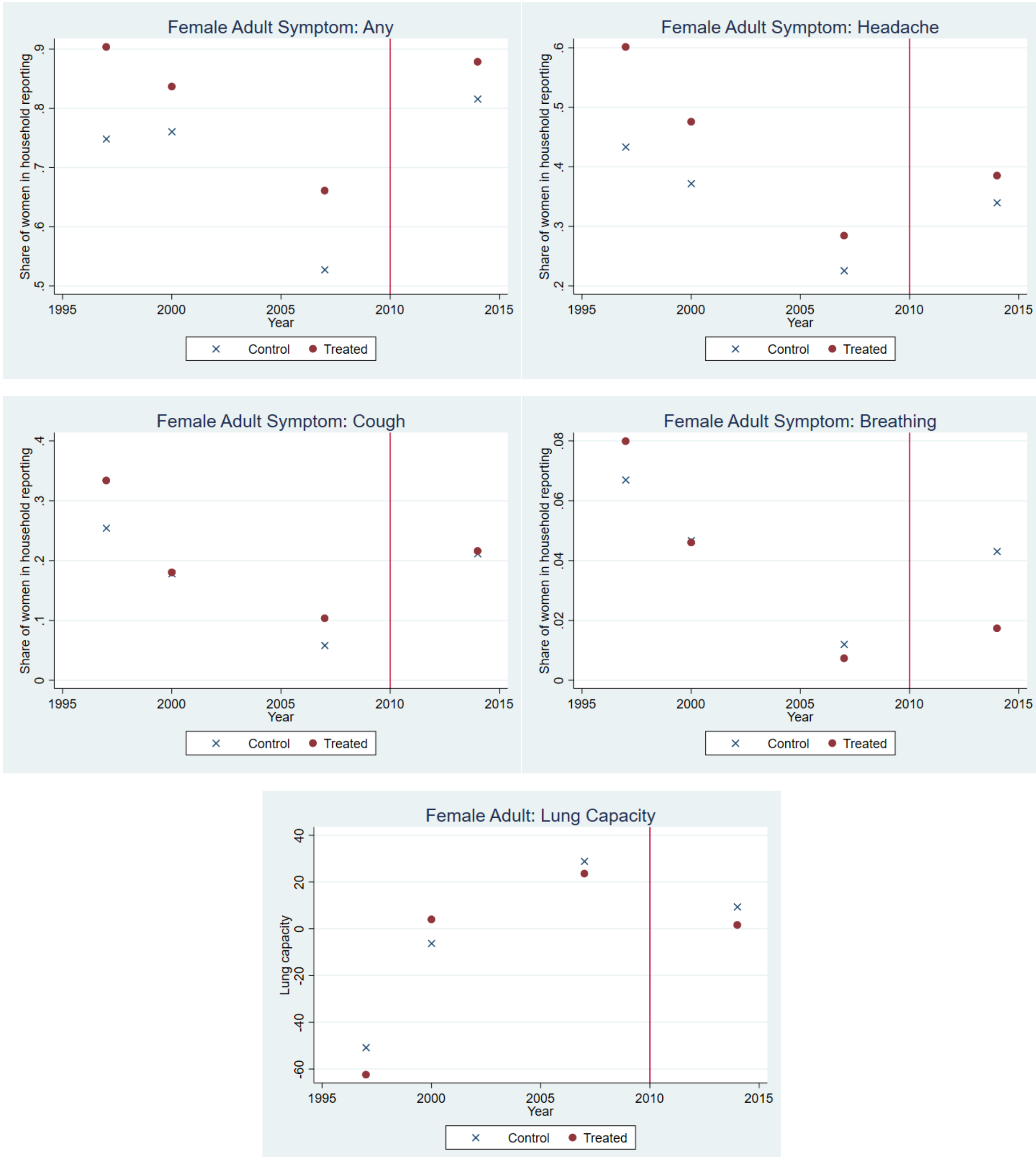
Notes: Treated indicates Lombok and control indicates Sumbawa. Data from 1997, 2000, and 2007 are pre-treatment. Data in 2014 is post-treatment. Red line indicates time of treatment in Lombok in 2010.

For health outcomes, Figure 35 presents the averages for women's self-reported health status and lung capacity across the 4 waves of the survey in Lombok and Sumbawa. The share of women within households reporting any illness ranges between 50 and 90%, with self-reports being slightly higher in Lombok. A similar pattern is found for headache illness, which with the share of women reporting ranging between 30 and 60%. Slightly fewer households report cough illnesses in the weeks preceding the survey, and less than 1% report breathing illnesses. For both of these metrics, rates across Lombok and Sumbawa appear very similar over time, with trends moving downwards from Wave 3 to Wave 4 and upwards from Wave 4 to Wave 5. As the data asks about self-reported information in the 4 weeks prior to the survey, other factors such as weather may explain the changes from wave to wave (Verma and Imelda, 2023). Despite this limitation, importantly for all of the self-reported variables, parallel trends are displayed across the two study areas before the Conversion Program was implemented in Lombok. For both the age-corrected lung capacities, scores measured through breathing tests appear to be increasing over time, with levels and trends similar in both districts across all survey years.

Figure 36 presents the trends for children under 15[59]. For this sub-group, the data displays similar patterns and trends for women, except for self-reported headache illness, which is less common for children. Parallel trends are observed for all self-reported data as well as the biometric lung capacity before the program. As with women, self-reports of breathing illness appear to increase sharply in untreated Sumbawa after the Conversion Program. However, response rates are similarly low around 1%.

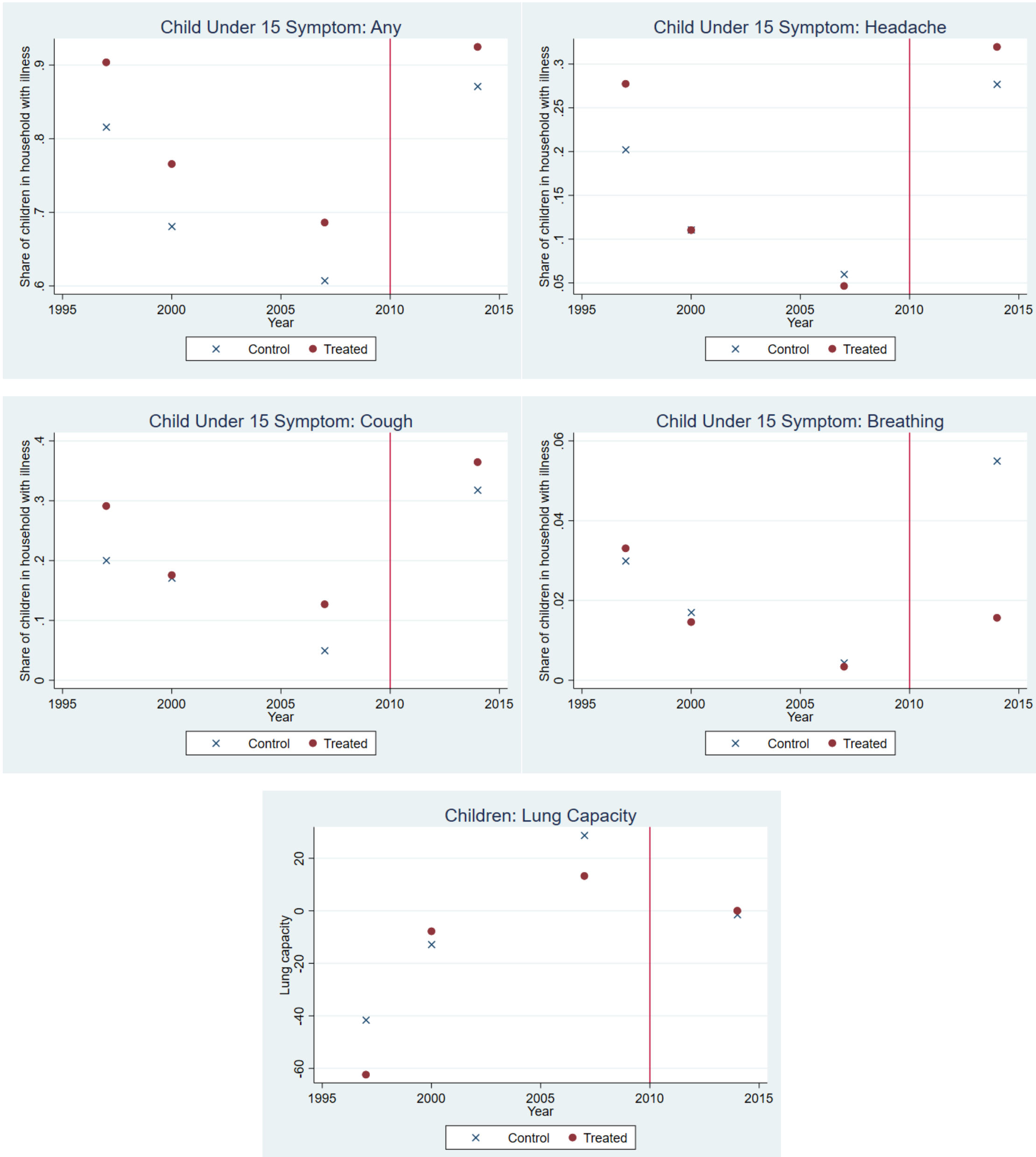
Figure 35. Outcome measures representing the health of women, across 4 waves of the survey.

[59] Trends and results for children under 5 are similar, so I present the data for children under 15.



Notes: Symptom data is from self-reports, while lung capacity is measured using a device with a trained health worker and is represented in mg/L. Treated indicates Lombok and control indicates Sumbawa. Data from 1997, 2000, and 2007 are pre-treatment. Data in 2014 is post-treatment. Red line indicates time of treatment in Lombok in 2010.

Figure 36. Outcome measures representing the health of children under 15, across 4 waves of the survey.



Notes: Treated indicates Lombok and control indicates Sumbawa. Data from 1997, 2000, and 2007 are pre-treatment. Data in 2014 is post-treatment. Red line indicates time of treatment in Lombok in 2010.

In addition to data on LPG use and the outcomes of interest, Table 37 provides descriptive statistics on the socio-economic control variables, examining trends between Wave 3 (2000) and Wave 4 (2007). Overall, the changes over time in these variables tend to be similar across households in Lombok and Sumbawa. In terms of household size, households in both areas are composed of around 6.5 members in 2007, and trends between Wave 3 (2000) and Wave 4 (2007) are similar. The same trends hold for per capita expenditure, although expenditures are slightly higher in Lombok. Electricity access levels in 2007 are high in both areas; however, Sumbawa experienced growth in electricity access from 2000 to 2007, and as a result the differences are statistically significant. In terms of household head, both districts are primary patriarchal and with trends not changing much over time. The same holds for schooling of household head, although levels of female head schooling are lower in both districts. In terms of financial inclusion, levels are low in both districts (at around 20% of households reporting holding a savings account), with trends not changing between 2000 and 2007. The final variable in Table 37 represents the second language spoken at home, after Indonesian, to proxy for culture. Households in Lombok disproportionally speak Sasak compared to households in Sumbawa, with the change statistically significant over time.

Table 37. Averages and trends in observed variables among treated (Lombok) and control (Sumbawa) districts.

	Average in 2007		Percentage change 2000-2007		
Variable	**Treated (Lombok)**	**Control (Sumbawa)**	**Treated (Lombok)**	**Control (Sumbawa)**	**T-stat of difference**
Household size	6.40	6.61	0.07	0.06	-0.55
Per capita expenditure (IDR)	398,399	319,516	1.97	1.80	-0.33
Electricity access	0.88	0.92	0.03	0.16	**3.11**
Household head male	0.68	0.82	0.11	0.10	-0.44
Male head attend school	0.81	0.92	0.02	0.04	0.65
Female head attend school	0.57	0.72	0.01	0.07	1.52
Savings account	0.18	0.17	-0.05	-0.05	0.02
Second language Sasak	0.97	0.13	0.09	0.00	**-3.87**

Notes: Bold in the final column represents statistically significant differences at the 5% level. IDR indicates Indonesia Rupiah.

The ability of the DID estimation to warrant a causal interpretation of the Conversion Program on household outcomes rests on three identifying assumptions: (1) parallel trends in the outcomes of interest among treated and untreated households before the program, (2) balance in observable characteristics, and (3) treated households do not interact with untreated households, or engage in reporting bias. Each identifying assumption is addressed in turn below.

The first and most important identifying assumption is that the outcome variables display similar trends before the program started. In other words, the intuition is that the trends in outcomes would be the same in both districts in the absence of the treatment, and that the treatment induces a deviation from this common trend (Angrist and Pischke, 2008). The existence of 3 pre-treatment waves and the descriptive results presented in Figure 33, Figure 34, Figure 35, and Figure 36 show evidence of parallel trends for fuel expenditures, education expenditures, self-reported health symptoms, and the lung capacity metric.

Second, if observable characteristics are balanced among treated and untreated districts, this lends credibility to the assumption of parallel trends in the absence of treatment and thus the causal interpretation of the estimates (Angrist and Pischke, 2008; Galiani and Schargrodsky, 2010). Observing similar trends makes it less likely that unobserved factors are driving the relationship between the treatment and expenditure and health outcomes. The balancing table presented in Table 37 suggests in general that the socio-economic trends (and often, the levels) in Lombok and Sumbawa are similar. While differences in electricity access trends between 2000 and 2007 exist, this is unlikely to confound the relationship of interest for two reasons. First, Lombok exhibited high levels of electricity access in 2000 of 85%, so mechanically the percentage change is lower compared to Sumbawa which was at 76% in 2000. The two districts appear to be converging: as of 2007, electricity access in both districts was very similar, around 90%. Second, as of 2014, electric stoves were virtually non-existent in the province, so electricity access is unlikely to affect decision of which fuel and stove to use. While electricity access may be correlated with the availability of other energy sources (e.g. urban areas may have more electricity access but have less accessible fuelwood), both districts are in rural settings. Furthermore, while the second language spoken is different in both districts, this is an imperfect measure for cultural differences (it was the best available proxy from the survey), and may not represent distinct cultural patterns which affect the relationship of interest. Nevertheless, these differences are controlled for in the empirical analysis.

Third, for the DID to produce reliable estimates of the *ITT*, one assumption is the stable unit treatment value assumption, or SUTVA. Under this assumption, the actions taken by households in the treated district of Lombok are independent of the actions taken by households in the control district of Sumbawa. Given the proximity between the two districts, there might exist spillovers and general equilibrium effects. For instance, LPG demand may increase in Lombok due to the program, which may increase prices and thus reduce LPG use in Sumbawa, confounding the relationship of interest. Another possibility is that households in Lombok experience the benefits of LPG use and communicate these benefits to their friends and relatives in Sumbawa, indirectly increasing LPG adoption in Sumbawa.

5.4. Results

5.4.1. Impacts on LPG use and expenditures

Table 38 shows the impact of the Conversion Program on take-up rates of LPG. Consistent with the descriptive statistics shown in Table 36 and Figure 32, the Conversion Program's introduction in Lombok led to a large uptake of LPG by around 54% compared to households in Sumbawa who were not included in the program. This figure suggests a strong and sustained take-up rate, as households in Lombok report using LPG as their main stove almost four years after the intervention took place in 2010.

Table 38. Impact of the conversion program on LPG use

	(1) LPG use
Conversion Program	.540***
	(.034)
Observations	804
Number of HH	402
Household FE	YES
Year FE	YES

Notes: Households in Lombok (239) are the treated group and households in Sumbawa (163) are the control group. Robust standard errors are included. * p < 0.10, ** p < 0.05, *** p < 0.01.

Table 39 shows that the Conversion Program led to a sharp decline in fuel expenditures but did not significantly impact other expenditure categories, such as education[60]. On fuel expenditures, the DID model finds annual fuel savings of around IDR 13,000 (around 13% of the mean and around 1% of total expenditures) which is statistically significant at the 1% level. This ITT estimate can be scaled-up to produce a local average treatment effect (LATE), if it is scaled-up by take-up proportion. Take-up of LPG is estimated at around 55% (Table 38), and if this is multiplied by the ITT, the LATE provides a fuel savings estimate of around IDR 7,250. While education expenditures similarly decrease for the treated group of households in Lombok, the point estimate is not

[60] Following (BenYishay *et al.*, 2020), the outcome data are top-end windsorized at the 95th percentile to account for extremely large values provided in the surveys.

statistically significant. This lack of significant finding is also shown for other expenditure categories, including food, housing, and health.

Table 39. Impacts of the Conversion Program on fuel and education expenditures

	(1) Fuel	(2) Education
Conversion Program	-13,132***	-34,485
	(3,487)	(26,773)
Observations	804	804
Number of HH	402	402
Household FE	YES	YES
Year FE	YES	YES
Outliers	W95	W95

Notes: Households in Lombok (239) are the treated group and households in Sumbawa (163) are the control group. Robust standard errors are included. * p < 0.10, ** p < 0.05, *** p < 0.01. Expenditure figures are in Indonesian Rupiah.

The strong impact on fuel expenditures is also explored in more detail in Table 25. As the Conversion Program was free, it greatly reduced the cost of transition to cleaner cooking for households, and we might expect poor households to have benefitted more (Budya and Arofat, 2011). Households are categorized as poor or non-poor based on their total expenditure, with those in the bottom 20% considered poor and those in the top 80% as non-poor, following guidance from the literature (Bacon, Bhattacharya and Kojima, 2010; World Bank, 2023b). Evidence is found that the Conversion Program's impacts were progressive, as poor households reduced fuel expenditures by around 28,000 IDR (Column 1), compared to non-poor households who benefited by 13,000 (Column 2). This two-fold decrease in fuel expenditures may indicate poor households are less likely to "fuel-stack" and use other, more costly fuels in addition to subsidized LPG.

Table 40. Heterogenous impacts on fuel expenditures for poor and non-poor households

	(1) Fuel expenditures – Poor	(2) Fuel expenditures – Nonpoor
Conversion Program	-28,801***	-12,960***
	(4,728)	(3,795)
Obs	804	804
N of HH	402	402
HH FE	Yes	Yes

Time FE	Yes	Yes

Note: Households in Lombok (239) are the treated group and households in Sumbawa (163) are the control group. Robust standard errors are included. * p < 0.10, ** p < 0.05, *** p < 0.01. Expenditure figures are in Indonesian Rupiah. Poor households are calculated as those in the bottom 20% of the consumption distribution, while non-poor households are those in the top 80%.

5.4.2. Impacts on health

The results exploring the impacts of the Conversion Program on health outcomes for women and children are presented in Table 41, Table 42, Table 43, Table 44, and Table 45. In terms of self-reported health outcomes, both women and children in treated households do not appear to experience any change in reports of any illness (Table 41) or for headache illness (Table 42). However, there is evidence that women in treated households do report fewer illnesses related to both cough (Table 43) and breathing (Table 44), which are statistically significant at the 10% and 5% levels, respectively. The point estimates are equivalent to around a 10% reduction in reporting either illness in the four weeks preceding the survey. Children under 15 in treated households do not experience the same benefits, with the point estimates close to zero for both cough and breathing.

Table 45 examines the impacts of the Conversion Program on age-corrected lung capacity, measured in mg/L. For both women and children, this biometric indicator remains unchanged when comparing treated households in Lombok to their counterparts in Sumbawa. The sample sizes drop when examining lung capacity, as measurements are missing in some of the households with self-reported data.

Table 41. Impacts on the share of women and children under 15 within each household reporting any illness.

	(1)	(2)
Regressor	Women	Children
Conversion Program	-0.078	0.068
	(0.057)	(0.090)
Observations	804	507
Number of HH	402	318
Household FE	YES	YES
Year FE	YES	YES

Notes: Figures represent the share of women and children (under 15) within each household reporting any illness in the 4 weeks prior to the survey. Robust standard errors are included. * p < 0.10, ** p < 0.05, *** p < 0.01*

Table 42. Impacts on the share of women and children under 15 within each household reporting headache illness.

	(1)	(2)
Regressor	Women	Children
Conversion Program	-0.046	-0.018
	(0.063)	(0.088)
Observations	804	507
Number of HH	402	318
Household FE	YES	YES
Year FE	YES	YES

Notes: Figures represent the share of women and children (under 15) within each household reporting headache illness in the 4 weeks prior to the survey. Robust standard errors are included. * p < 0.10, ** p < 0.05, *** p < 0.01*

Table 43. Impacts on the share of women and children under 15 within each household reporting cough illness.

	(1)	(2)
Regressor	Women	Children
Conversion Program	-0.119*	0.031
	(0.063)	(0.095)
Observations	804	507
Number of HH	402	318
Household FE	YES	YES
Year FE	YES	YES

Notes: Figures represent the share of women and children (under 15) within each household reporting cough illness in the 4 weeks prior to the survey. Robust standard errors are included. * p < 0.10, ** p < 0.05, *** p < 0.01*

Table 44. Impacts on the share of women and children under 15 within each household reporting breathing illness.

	(1)	(2)
Regressor	Women	Children
Conversion Program	-0.071**	-0.041
	(0.036)	(0.036)
Observations	804	507
Number of HH	402	318
Household FE	YES	YES

Year FE	YES	YES

Notes: Figures represent the share of women and children (under 15) within each household reporting breathing illness in the 4 weeks prior to the survey. Robust standard errors are included. * p < 0.10, ** p < 0.05, *** p < 0.01*

Table 45. Impacts on age-corrected lung capacity for women and children under 15 within each household.

	(1)	(2)
Regressor	Women	Children
Conversion Program	-3.64	-2.89
	(7.16)	(16.74)
Observations	769	325
Number of HH	395	251
Household FE	YES	YES
Year FE	YES	YES

Notes: Figures represent age-corrected lung capacity for women and the average for children in each household. Lung capacity is measured in mg/L and is collected by a trained surveyor. Robust standard errors are included. * p < 0.10, ** p < 0.05, *** p < 0.01*

5.4.3. Comparison with prior studies

The results from this paper are relatively similar to previous quasi-experimental studies regarding fuel and overall expenditures (Imelda, 2018) but differ on health outcomes (Imelda, 2020; Verma and Imelda, 2023). In this paper, I find a reduction of monthly fuel expenditures by around 13,000 INR (approximately USD $1.00), which accounts to a 13% savings and is statistically significant at the 1% level. This finding is consistent albeit a lower estimate than previous work with government surveys in urban areas suggesting a 30% reduction in fuel expenditures (Budya and Arofat, 2011; Andadari, Mulder and Rietveld, 2014) and quasi-experimental work across the country finding a 40% reduction among early-treated households (Imelda, 2018). The point estimates from this paper of monthly fuel savings of USD $1.02 is slightly lower in range to surveys conducted with small samples (up to USD $1.64 in (Budya and Arofat, 2011; Andadari, Mulder and Rietveld, 2014) and similar to the point estimate from the working paper from Imelda (2018) of USD $1.19, which is significant at the 5% level.

On how households reallocate fuel savings, the findings from this paper are in line with the results from (Imelda, 2018). Both sets of results do not find any statistically significant reallocation towards education, food, housing, health, or other consumption categories. Given that fuel savings accounted for 1-2% of total expenditures for households in Lombok (this paper) as well as for households across the country (Imelda, 2018), this is unsurprising, especially for education expenditures which are heavily subsidized across Indonesia (World Bank, 2023b).

However, the results on health outcomes differ when compared to prior quasi-experimental work focusing on morbidity (Verma and Imelda, 2023). For the sample in West Nusa Tenggara, I find evidence that women report fewer illnesses related to indoor air pollution but find no change in the biometric indicator of lung capacity. Published work focusing on the entire country (Verma and Imelda, 2023) finds the opposite with no significant change for self-reported health outcomes for women, but a 4% and statistically meaningful improvement in lung capacity.

The diverging results between this paper and previous quasi-experimental work on Indonesia's Conversion Program may be due to several factors, with differences in (1) treated and control groups, (2), measurement of health outcomes, and (3) policy implementation data sources. Each factor is discussed in turn.

5.4.3.1. Treated and control groups

For the paper on fuel expenditures, Imelda (2018) examine program impacts across the country and classify treated households who reside in districts that receive the program before 2011, with the control group as households who live in districts that get the program after 2011. For the paper on health by (Verma and Imelda, 2023), households who received LPG in Phase 1 (2007-08) were excluded from the program, and the sample is restricted to individuals in households who did not report using LPG in the 2007 wave of the IFLS. As a result, the findings of these papers can be interpreted as the impact of the duration of the program (with treated households experiencing the program for, on average, 3 more years). This paper compares households who were treated for 2-3 years (in Lombok) compared to households who were never treated (in Sumbawa).

Nevertheless, the results on fuel expenditure savings found in this paper (1.02 USD per month) are very similar to the previous study (Imelda, 2018) (1.19 USD per month). One might expect households to be able to accrue fuel expenditure savings soon after receiving the LPG stove and starter kit, as suggested from government follow-up surveys (Pertamina, 2012; Andadari, Mulder and Rietveld, 2014). For this reason, comparing households treated for 4 years to households treated for 1 year may provide a lower estimate (as in (Imelda, 2018)) than comparing households treated for 2-3 years to those not treated by the program at all (this paper). Along these lines, later-treated areas may have anticipated the transition to LPG and may have started to use it even before the program was officially implemented in their area. In this paper, anticipation is unlikely as Sumbawa was not treated in the program at the point of data collection in 2014. Alternatively, it may take time for households to become familiar with and learn how to efficiently use the new LPG stove. If the learning period is longer than a year, then the results from (Imelda, 2018) and this paper are more directly comparable.

For health outcomes, when examining the results of Verma and Imelda (2023) closely, more consistency is found with the evidence found in this paper. The headline result that women experience a statistically significant improvement in lung capacity by 4% is when comparing households treated for 5-6 years compared to households treated for 0-2 years. However, when comparing households treated for 3-4 years compared to those treated for 0-2 years (Figure A7 in the Online Appendix of (Verma and Imelda, 2023)), the results are not statistically different from zero. This second result is more comparable to the setting of this paper and is consistent with the lack of improvement in lung capacity for women in the sample of West Nusa Tenggara. One explanation is that it may take time – in this context, at least 5 years – for LPG stove use to lead to improvements in health for women, in line with evidence from the epidemiological literature (Chillrud *et al.*, 2021).

Another main difference between this paper and prior work is the sample. While previous work (Imelda, 2018; Verma and Imelda, 2023) examines the entire country (~83% of Indonesia is covered in the IFLS), this paper only focuses on households within the province of West Nusa Tenggara. In Imelda (2018), the classification using 2011 as a cutoff leads to a skewed distribution of treated and control groups, with 16,226 households in the treated group and 3,868 in the control group. When omitting households who received LPG in Phase 1, the sample size of (Verma and Imelda, 2023) consists of 3,816 individuals in the treated group, and 2,255 in the control group. In terms of identification, the second study makes a convincing argument showing similar pre-trends in health outcomes for treated and control households, suggesting a valid counterfactual (Table B2 in the Online Appendix of (Verma and Imelda, 2023)). For the paper on consumption (Imelda, 2018), while parallel trends on utility bills and the quantity of kerosene consumed are shown, the test on fuel expenditures is not displayed (fuel expenditures account for 40% of overall utility bills). Similarly, parallel trends on monthly expenditures do not hold (Table 3.1 in (Imelda, 2018)), which opens up the possibility the control group may not precisely estimate what would have happened to the treatment group in the absence of the program. For instance, if households in the control (later-treated) group were part of cash transfer schemes that are common in Indonesia (World Bank, 2023b), they may be less concerned about fuel savings, which would bias the estimates upwards.

5.4.3.2. Measurement of health outcomes

While the use of the expenditure data (and the results found) is similar between this study and (Imelda, 2018), the interpretation of the health data in the IFLS are not straightforward and might explain the differences found. Verma and Imelda (2023) argue that lung capacity may be a better metric of overall respiratory health compared to self-reported symptoms which may be prone to measurement error, systematic bias (Bertrand and Mullainathan, 2001), and can be impacted by contemporaneous weather conditions. For example, if the month of survey implementation changed from wave to wave, or if there was a lack of rain or wind in the weeks before the surveyor asked questions to the household, the changes in self-reported health outcomes of cough or breathing for women found in this paper may be due to these factors instead of from the Conversion Program and is a threat to identification. On the other hand, lung capacity measurements might also be prone to measurement error, if the devices used to collect the biometric data improved over time, or if training to health workers implementing the survey changed from wave to wave. While the difference in results may be attributable to program impacts across the country compared to West Nusa Tenggara, it is notable that both papers do not find substantially large improvements in self-reported health status or lung capacity.

These results of small to moderate health improvements from this paper and from Verma and Imelda (2023) can also be fit into a broader frame examining health impacts of clean cooking. Hanna et al. (2016), the paper on India, finds a similar result of modest if any health improvements, but attribute this to the lack of use of the improved cookstoves in years after the program for treated households. In Indonesia, as evidenced from this paper and from Verma and Imelda (2023), cookstove adoption remains high years after implementation; as a result, this explanation cannot suffice in this context. However, an alternative explanation that might explain the relatively moderate health improvements in Indonesia is that treated households, due to savings on cooking fuel from moving from kerosene to LPG, spent these resources on lighting fuel, which is typically from kerosene or biomass. Increased use of kerosene or biomass for lighting might offset the health benefits from improved cookstoves use (Barron and Torero, 2017). More generally, households may be living in highly polluted environments to begin with, and any marginal change may not lead to sustained and large improvements in lung capacity or self-reported health status (Chillrud *et al.*, 2021). Future research may want to consider collecting detailed data on pollution exposure, as well as cooking and lighting fuel expenditures which are currently unavailable in the IFLS.

5.4.3.3. Policy implementation data sources

Another difference between this paper and prior quasi-experimental work pertains to how implementation of the LPG Conversion Program is tracked and included in the analysis. This paper uses publicly available information sources from the policy literature to understand program implementation in West Nusa Tenggara and identify the pause in roll-out that resulted in Lombok receiving the program in 2011 and Sumbawa not receiving it at the point of data collection in 2014. The other two studies (Imelda, 2018; Verma and Imelda, 2023) rely on restricted administrative data on the program roll-out. While details on this restricted data are unavailable, it is possible that this dataset provided more accurate implementation data than the publicly available maps provided in Pertamina (2012) and Wiratmaja (2016). However, if the restricted administrative dataset provided similar information to early policy documents and did not account for changes in policy implementation that occurred, the “later-treated” districts in (Imelda, 2018) and Verma and Imelda (2023) may actually have been districts that were not at all treated by the program. If this is the case, the empirical setup and methodology exhibit more similarity to this paper and the results more directly comparable.

5.5. Robustness and limitations

In this section I examine the robustness of the above findings and discuss several limitations with the study.

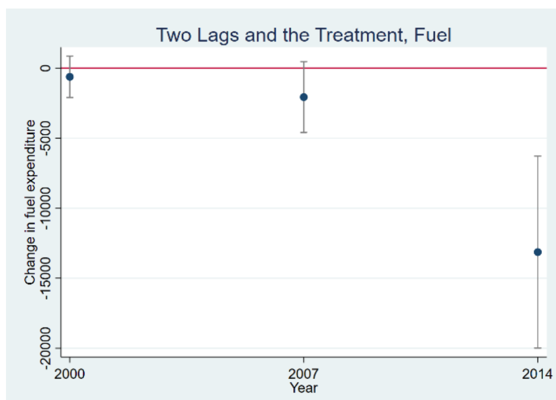
5.5.1. Robustness checks

5.5.1.1. Placebo lag tests for changes in fuel expenditures and self-reported health outcomes

The DID estimation strategy found a statistically significant drop in fuel expenditures, women’s self-reported cough illness, and women’s self-reported breathing illness. To further examine these findings, I run placebo lag tests using previous survey waves to explore if these three outcomes changed before the program was implemented. If I find that fuel expenditures, cough illness, and breathing illness changed between Wave 2 and Wave 3, or Wave 3 or Wave 4 (before the Conversion Program was implemented), then this reduces the confidence that the program led to a causal impact. In addition to lagged placebos, leads are often estimated too, to examine the change in periods after the treatment as in Hainmueller & Hangartner (2019); however, in this circumstance there is no survey after Wave 5 to examine leads.

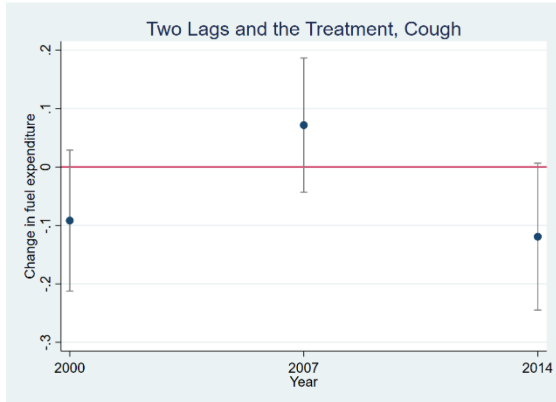
The results for fuel expenditures are presented in Figure 37. There is strong evidence that the change in fuel expenditures was near zero and highly noisy when comparing Wave 2 to 3 (indicated as 2000 in the figure) as well as from Wave 3 to 4 (indicated as 2007). This changes drastically when examining the difference between Wave 4 and Wave 5 (which shows the changes between 2007 and 2014). The consistency of these findings increase the confidence that the Conversion Program had a causal impact on fuel expenditures. The results for self-reported cough and breathing illness is reported in Figure 38 and Figure 39. All lagged placebos are not statistically significant at the 5% level, which adds confidence that the reductions in self-reported cough and breathing illnesses found in this paper can be attributed to impacts of the Conversion Program rather than spurious factors.

Figure 37. Placebo lagged impacts (2000 and 2007) of the conversion program, and actual impact of the conversion program (2014) on fuel expenditures.



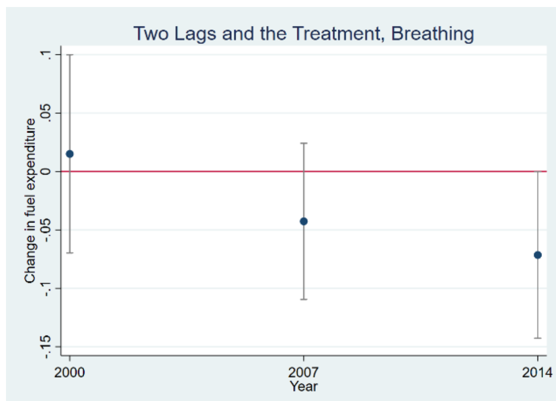
Notes: Vertical lines indicate 95% confidence intervals.

Figure 38. Placebo lagged impacts (2000 and 2007) of the conversion program, and actual impact of the conversion program (2014) on cough illness for women.



Notes: Vertical lines indicate 95% confidence intervals.

Figure 39. Placebo lagged impacts (2000 and 2007) of the conversion program, and actual impact of the conversion program (2014) on breathing illness for women.



Notes: Vertical lines indicate 95% confidence intervals.

5.5.1.2. Larger sample

The estimations provided above are run on a sample of 402 households who were surveyed in all Waves from 1997 to 2014, with the focus on the differences between 2007 and 2014. However, as the survey evolved, more households were added. In this section, I extend the analysis to include to households who were surveyed in 2007 and 2014, regardless of whether they were included before. This "larger" sample includes 548 households (298 in the treated district of Lombok and 250 in untreated Sumbawa).

The full results are presented in Appendix E1. A comparison of the results when using this "larger" sample to the "smaller" sample presented above both find statistically significant and large savings on fuel expenditures, but do reveal inconsistencies for the other measures. For education expenditures, while the small sample found no impact, including the larger sample reveals an increase in expenditures of around 46,000 INR for treated households that is statistically significant at the 5% level. Similarly, for women's health outcomes, the coefficients for any illness, breathing illness, and cough illness that were previously significant for the small sample are no longer significant with the larger sample. Outcomes for children remain small and insignificant for both samples.

One potential explanation to reveal the inconsistent results is due to the larger statistical power when using a larger sample of households. Another reason might be that the households added to the survey in later years (e.g. the 146 households not present before 2007) were systematically different from the 402 households in the smaller sample. However, this appears unlikely since I find that new households are similar to existing households across a number of socio-economic characteristics, both in Lombok and Sumbawa. Nevertheless, the lack of consistency in the health results reduces the confidence in the significant results found under some models for women's health when using the small sample. However, similar to the placebo lag tests, the consistent results for fuel expenditure increase the confidence that the Conversion Program did indeed reduce fuel expenditures for treated households.

5.5.1.3. Male health outcomes

As discussed in the literature review, the improvements in health as a result of moving toward a cleaner cooking fuel are typically concentrated in the sub-groups of women and children. This section examines health outcomes for males, which are unlikely to be affected by a change in cooking patterns, since males are much less exposed to the indoor air pollution since they spend much less time in the household (Hanna, Duflo and Greenstone, 2016).

The results for male health outcomes are provided in Appendix E2, with a focus on the differences between health outcomes for women and men. Using self-reported data, no effects are found for any illness and headache illness for men, as well as lung capacity. However, breathing and cough illness for men does reduce in a statistically significant way, like the effect on women. This may reduce confidence that the Conversion Program is the mechanism leading to reduced illness, especially in a context where 98% of households have women as the main cook. It is possible that men do some of the cooking, or do experience spillover benefits within the household, which can be explored in future analyses with detailed time-use data.

5.5.2. Limitations

This paper has several limitations which are presented below, alongside a discussion of external validity.

First, the measurement of health outcomes and unsubsidized fuel expenditures may be inaccurate and imprecise. I focus on self-reported symptoms of illnesses in the 4 weeks prior to the survey, but health problems related to indoor air pollution might be chronic, which would not be reflected in this measure (K. R. Smith *et al.*, 2014). In addition, non-linearity and threshold effects may operate in this context, where indoor air pollution needs to be reduced below a certain threshold before delivering an apparent health outcome (Hanna, Duflo and Greenstone, 2016). For this paper, I do not observe pollution data at the household level, but advances in pollution monitoring are making household-level measurement a reality in many contexts (Somanathan *et al.*, 2022). Furthermore, household fuel expenditures are distorted by subsidy regimes for both kerosene and LPG, which hide the true costs of fuel use (Thoday *et al.*, 2018). The IFLS does not differentiate unsubsidized fuel expenditures from subsidized fuel expenditures in the data. Future work to collect high resolution data on both categories can help inform whether the finding from this paper is an under-estimate or an over-estimate of fuel savings.

Furthermore, an important issue related to the survey instrument is measurement error. I assume that households do not systematically provide answers that they believe the surveyors are looking for, a phenomenon known as 'reporting bias' (Bertrand and Mullainathan, 2001). One might argue that surveyed households might believe that IFLS surveyors were connected to the Conversion Program, and as a result reported having used the LPG stove to keep the product. Furthermore, households may over-report symptoms if they believe the survey is tied to support from NGOs or improved access to health facilities. However, this is unlikely, given the IFLS has been implemented since 1994 (15 years before the Conversion Program) and the questions on fuel use and self-reported health outcomes are a very small part of the survey instrument compared to all the questions asked.

Another caveat to consider when employing the differential timing of a government policy to assess the program's causal impact is anticipation (Bertrand, Duflo and Mullainathan, 2004). In such a case, areas that are treated at later points in time might anticipate treatment and thus change their behavior accordingly. For instance, it is possible that households in Sumbawa expected to receive a free LPG stove in the future, and as a result spent more money on fuel expenditures for kerosene before the program was implemented. However, anticipation is unlikely to have taken place. While households were aware of the Conversion Program, given its large scale nature, national coverage, and communication from the government, they were unaware as to when their district would get treated at which time (Andadari, Mulder and Rietveld, 2014). The first publicly available document that provided information on how the program would be timed was only released in 2013 (PT Pertamina, 2013). As a result, given a lack of knowledge of the exact timing of implementation, it is unlikely households in untreated areas anticipated receiving the stove.

One further limitation is related to the lack of detailed micro-data on program implementation and compliance. Around 40% of households in Lombok did not report using LPG as their main stove after the treatment. It is unclear whether these households received the treatment and were non-compliers, or did not receive the treatment at all. Future research and data collection can investigate whether the 40% not reporting LPG in Lombok were given the program and chose not to comply, or were not targeted by the program. Future research on policy implementation at the local level can help inform whether the *ITT* presented here could also be interpreted as the *ATT* and how similar the 40% not reporting LPG is to the 60% that do report it.

Additionally, the data presented in this paper has long gaps between collection. While other studies using IV or DID in the literature use annual data (e.g. (Hainmueller and Hangartner, 2019)), there is a 3-7 year gap between survey waves in the IFLS. One might argue that these large data gaps are the reason why I do not find significant and large impacts of the program on health outcomes of children, or for lung capacity. It is possible that the benefits accrued during the few years after the treatment but before the first post-treatment survey. Alternatively, given that health outcomes are a 'stock' rather than a 'flow', it may take time for health conditions to improve, justifying the longer period of analysis.

Finally, this paper does not consider "fuel-stacking", whereby households use multiple cooking modalities at the same time. I only know the main stove that households use for cooking, but I do not know the distribution. For example, a household who reports using LPG as the main stove may also use kerosene for 40% of their cooking, or not at all. Related to this, one possibility that was not examined in this paper was the potential for heterogeneous effects. For instance, results may differ for a household that switches from kerosene to LPG compared to one that switches from firewood to LPG. Recent research suggests kerosene is highly polluting, but the epidemiological consensus is that pollution from firewood is still worse (Epstein *et al.*, 2013; K. R. Smith *et al.*, 2014). In this context, an issue when evaluating heterogeneous effects is the small sample size. Of the 139 households that converted to LPG in Lombok, 48 converted from firewood and 91 from kerosene. Future data collection to better understand the full portfolio of options households use for cooking (as well as lighting) can be used to explore heterogenous effects and interactions.

5.5.2.1. External validity

A relevant question is to what extent the results found in this paper are valid in other parts of Indonesia, in other countries, or at different times. One might argue that external validity to other areas in Indonesia might be appropriate, given the program was also implemented in phases in other areas of the country. However, given Indonesia's island geography, each province is very different from each other in terms of energy prices, LPG infrastructure, and socio-economic outcomes (Rentschler and Kornejew, 2017), mechanisms through which the Conversion Program operate. Second, even within the same country, policy implementation can vary significantly in developing countries, due to differing levels of administrative capacity (De Ree *et al.*, 2018). As a result, the findings of this paper may not be generalizable even within the same country. This may explain the differing results on health outcomes compared to papers exploring the Conversion Program across the country (Imelda, 2020; Verma and Imelda, 2023).

In terms of external validity to other countries, this appears even more unlikely since clean cooking interventions vary in design based on the country that is implementing it (Toman and Bluffstone, 2017). Regarding generalizability across time, one would have to consider the political economy aspects of the program. Indonesia's Conversion Program was implemented in a time of high government subsidies for kerosene (and was the main driver for the program); as a result, the government had an incentive to implement the program as successfully as possible (Toft, Beaton and Lontoh, 2016). However, in times of less fiscal pressure, the Conversion Program may have been implemented differently. A final point regarding external validity is that similar cookstove programs are often implemented by NGOs instead of governments, which might reduce their impact if recipients perceive NGOs to be less credible (Toman and Bluffstone, 2017). Given these multiple concerns on external validity, one would have to be careful to extrapolate the results of this paper beyond the Indonesian province of West Nusa Tengarra.

5.6. Conclusion and policy implications

Indonesia embarked on a large-scale program to convert 50 million households to using cleaner LPG cookstoves starting in 2007. This paper contributes generally to the literature on fuel conversion across The Global South and Indonesia's program specifically to evaluate the causal impacts in West Nusa Tengarra province. Using a differences-in-differences estimation strategy and the delay in implementation of the program across districts, I find that the Conversion Program led to a sharp increase in LPG use by 55% and substantial fuel savings of around 13% for treated households which is robust across estimators and samples, with higher benefits accruing for poorer households. While these savings are economically meaningful, they account for less than 1% of total expenditures, and I do not find evidence that households reallocate consumption towards pro-social spending like education. In terms of health outcomes, I do find evidence that women included in the program do report fewer illnesses related to indoor air pollution, but this result is small in magnitude, may be prone to measurement error, and does not measure particulate exposure at the household level.

As the program was not primarily implemented to reduce household fuel expenditures or improve health, it is not surprising to find household benefits to be moderate in this context. Nevertheless, it is important to note broader achievements of the program. The driving force behind the policy was to reduce the ever-growing subsidy bill for kerosene in Indonesia and estimates suggest the program saved $15.6 billion USD within 10 years (Wiratmaja, 2016). Past reforms of the energy sector in Indonesia which produced fiscal savings were re-directed towards social protection programs for the poor, re-directing benefits for households progressively (Bacon, Bhattacharya and Kojima, 2010). Future research on the Conversion Program may want to investigate how these $15.6 billion USD in savings were reallocated and if households received any further monetary benefit in addition to the reduction in fuel expenditures found in this paper and in other work.

Additionally, the high-take up rates of LPG both in West Nusa Tenggara and across Indonesia can be considered a policy success. The literature assessing similar programs of clean cookstoves distribution finds that households generally do not use the improved stove one or two years after receiving it (Hanna et al., 2016). The finding from this paper that over 60% of households in Lombok still report using LPG as their main stove 4 years after implementation is an achievement in this regard. One potential explanation is that households are more likely to use clean cookstoves when the distributor is from the government (as in Indonesia), compared to an NGO (as in most other cases). An alternative explanation for the high take-up is that the introduction of LPG was paired with a removal of kerosene, reducing outside options for households (Verma and Imelda, 2023). Future research may want to prioritize high-frequency and detailed data collection on the range of options that households use for cooking to better understand household preferences, the phenomenon of fuel stacking, and relationships to expenditure patterns and indoor air pollution.

Inspired by Indonesia's program, a number of other countries including India, Mexico, Ghana, Peru, and Thailand have started to promote LPG (Toft, Beaton and Lontoh, 2016; Quinn *et al.*, 2018). To maximize household co-benefits in addition to fiscal savings, future programs may want to be implemented with an explicit focus of saving poor households money and reducing indoor air pollution. For instance, roll-out of LPG can be prioritized towards households using biomass for cooking and paired with education campaigns on the to better understand cooking preferences and outline the benefits of reducing indoor air pollution. Programs may want to use recent innovations in pollution monitoring at the household level to better capture information on who is exposed to particulate matter and how effective LPG can be to reduce exposure. More generally, in terms of monitoring and evaluation, countries exploring roll-outs of LPG may want to design and run pilot programs and include both a control and treatment group to track outcomes at the monthly or annual frequency. Given the large-scale nature of these planned programs, evidence on program impacts could provide valuable input on policy design before scale-up.

Conclusion

The five chapters of this PhD thesis have examined sustainable development challenges across The Global South focusing on a range of environmental risks at various scales in a number of contexts. This section examines the external validity of the evidence gathered, provides implications for policy, and suggests directions for future research.

When interpreting the results gathered from the five chapters, a relevant question is how generalizable, or how externally valid, the evidence is across other contexts. While Chapter 1 examines global evidence at the sub-national level and is likely to be externally valid, Chapters 2-5 focus on individual case studies at the country-level in Vietnam, Nigeria, and Indonesia. The results from these four chapters may be only pertinent to the countries studied and may not be generalizable across other settings. List (2020) provides four key areas when addressing external validity to examine if the results can be generalized to different people, situations, stimuli, and time periods. These four areas include selection, attrition, naturalness, and scaling, and the last four chapters of this thesis are examined together across these criteria.

First, selection examines how representative the studied group is compared to the underlying population (List, 2020). Chapter 2 on environmental risks and poverty in Vietnam and Chapter 4 on floods in Nigeria use nationally representative panel household survey data. The city-level analysis of Chapter 3 focuses on Ho Chi Minh City, a large metropolis in south-central Vietnam which is quickly-growing and currently home to 9 million people and generates a quarter of Vietnam's GDP (World Population Review, 2023). Furthermore, Ho Chi Minh City surrounds the Saigon River and is located close to the coast. The results which find an over-exposure of slum locations to floods in Ho Chi Minh City may only be generalizable to large and highly dense cities with high flood risk and without strong urban planning infrastructure, which characterizes many growing cities across The Global South (Kocornik-Mina *et al.*, 2020). Chapter 5 on LPG Conversion in Indonesia only focuses on households in West Nusa Tenggara province, a relatively rural island with a low population density and a large agricultural sector. The results which find fuel expenditure savings and no robust health improvements of an LPG switch may only be applicable to this context, and may not pertain to other more urban or integrated areas of Indonesia, or in similar settings in other countries.

Second, attrition explores the compliance rates of subjects and documents reasons for attrition and non-compliance (List, 2020). The household data used in Chapters 2 and 5 are based on large-scale longitudinal surveys, with low attrition rates and significant resources spent to track households over time. While Chapter 4 on Nigeria similarly has a relatively low attrition rate (of 9%), the nature of the study opens the possibility that households dropped out of the sample due to the shock and were not included in the post-flood wave of the survey. While the attrition rate is not correlated with either metric of flood exposure, a certain number of households who experienced the most devastating impacts from the flood may have moved and were not included in the post-flood wave. If this is the case, this the results which find 20% drop in crop production and a 40% decrease in crop value can be treated as a lower-bound estimate.

Third, naturalness refers to the setting and timeframe of the study and relates mainly to experiments conducted by researchers (List, 2020). Chapters 2 and 3 rely on trends of environmental risk and household changes across time rather than specific shocks. The flood in Nigeria in 2012 that is the subject of Chapter 4, while "natural", was unprecedented in its scale. For this reason, the results from this study may be more generalizable to the impacts of extremely large floods in contexts of under-preparation by the government and the general population, which also unfortunately reflects many flood-prone areas across The Global South (Hallegatte *et al.*, 2020). Chapter 5 on LPG conversion in Indonesia is the closest to a "natural experiment". In this case, the intervention provided households with new cooking technology. While it was the first program of its kind across Indonesia, such programs are commonplace across The Global South (Krishnapriya *et al.*, 2021). Additionally, while the program in Indonesia may have been an "unnatural" shock at the start of implementation in 2007, Chapter 5 examines impacts years later during later phases of the program, where some level of familiarity with the program may have been established.

Finally, scaling requires an understanding of program effects across subsets of the general population and guidance on implementation (List, 2020). Chapter 5 describes Indonesia's program as being implemented first and foremost for fiscal purposes, with potential socio-economic and health benefits for households not prioritized explicitly during implementation. Several other countries across The Global South have started to design similar programs to promote the adoption of LPG. Policymakers in these countries may want to prioritize household benefits as part of the implementation for example by pairing LPG stove distribution with technologies to reduce indoor air pollution, to maximize potential co-benefits and increase the project's cost-benefit analysis. While Chapters 2-4 relate to natural phenomenon rather than program impacts, social protection schemes to support households prepare for weather shocks and recover from them were identified as policy priorities. Such schemes may want to be targeted to cover populations highly exposed to floods and environmental risks as well as flexible enough to be quickly scaled-up whenever a weather shock materializes.

With these caveats in mind in terms of external validity, the body of research from this thesis can inform global progress towards reaching the targets outlined in the SDGs. Chapter 1 on land and poverty provides global evidence that improving life on land (SDG 15) can itself support the livelihoods of households and reduce extreme poverty (SDG 1). While recent trends show a net reduction of tree cover globally by 2.4% since 2000 (Global Forest Watch, 2023), efforts to slow these rates, especially in low-income countries, can support progress in reaching SDGs 1 and 15. Chapter 2 on environmental risks in Vietnam shows how the relationship between households and the environment is context specific, with certain groups more impacted than others. Efforts to improve environmental quality or reduce air pollution can also support progress in achieving SDG 10 on reduced inequalities. Chapters 3 and 4 on floods in Vietnam and Nigeria show that poor people are highly exposed and vulnerable to floods, and taking climate action (SDG13) to better prepare for future flood events can help achieve zero hunger (SDG2) across The Global South. Finally, Chapter 5 on the LPG program in Indonesia provides an example of a country that has used a fiscal policy to promote a large-scale energy transition which may provide co-benefits in terms of affordable and clean energy (SDG7), improved health and well-being (SDG3) and gender equality (SDG5). Similar policies are being carried out across The Global South and may want to be designed in a way to maximize societal co-benefits.

The research from this thesis also provides several implications for policy-making across The Global South. One overarching implication is that there is an opportunity to explicitly embed environmental risk and climate considerations in general poverty reduction policies. For example, efforts to improve soil and land quality themselves lead to improvements in poverty, so government actions to incentivize smallholder households to reforest and protect landscapes can provide both economic and environmental benefits at the local level. Additionally, as extreme weather events can roll back progress on poverty reduction (Hallegatte, Bangalore, *et al.*, 2016), expanding existing social protection schemes to support livelihoods in the aftermath of a disaster can hasten recovery. Efforts such as Ethiopia's Productive Safety Net Program which provide work opportunities for households on projects that strengthen climate resilience can be built upon and tailored to each individual country context (Tenzing, 2020).

Water management is another area of critical importance for governments across The Global South. On one hand, irrigation has the potential to break the link between poor weather and adverse household outcomes and can be prioritized where feasible. Flood defenses and other hard infrastructure may want to be planned with climate change considerations in mind, instead of being benchmarked to historical water levels which may lock-in vulnerability. Additionally, while historically flood infrastructure has been directed towards areas with high asset values (with economic damages prioritized), policymakers may also want to include equity considerations and undertake projects in areas that protect households that are highly vulnerable to the impacts of floods.

More generally, policymakers may want to plan climate-smart cities of the future and consider rural livelihoods in this process. With ever-growing populations in urban areas across The Global South, leaders may want to encourage settlement in areas that are less prone to environmental risk and climate impacts. For example, opportunities to identify areas of the city less prone to extreme weather using satellite data, promoting development in these areas, and pairing these investments with public transit may allow households to reside in safer parts of the city but also access job opportunities. Additionally, the urban agenda is intertwined with rural livelihoods, as shocks in rural areas are a main reason why households migrate to cities (Selod and Shilpi, 2021). Strengthening rural livelihoods may reduce the migration pressure and help urban leaders better manage development in a climate-smart way. Overall, one implication from this thesis is that consulting communities in risk management strategies can be prioritized, to ensure households benefit from any intervention and are not made worse off.

This PhD thesis also outlines priorities for future research on data and methods. The chapters of this thesis has benefited greatly from advances in remote sensing and household survey data collection over the last decade. Areas of further improvement on data collection may want to be prioritized. In particular, linking data from global models and satellites with on the ground information can vastly improve our understanding of how environmental risks impact households locally. For example, to better measure flood impacts, databases on flood protection across The Global South can be better outlined and paired with new questions in household surveys asking farmers about plot-level strategies to protect their land from water intrusion. Household surveys may also want to measure household welfare beyond standard income and consumption metrics, to better understand detrimental health impacts or quality of life changes due to poor environmental conditions. Methodologically, future work can better integrate geo-spatial information with household surveys to conduct analyses while addressing uncertainties both from the satellite data (e.g. accounting for cloud cover) and on the precise location of the household (e.g. by randomly re-sampling household locations within the buffer of displacement).

In 2023, we lie halfway between the enactment of the SDGs in 2015 and their planned conclusion in 2030. Environmental risks and climate change pose a challenge to the implementation success across all 17 goals, as evidenced by the devastating floods during the rainy season of 2022 across The Global South in West Africa and South Asia (Al Jazeera, 2022; Hersher, 2022). The evidence gathered through the five chapters of this PhD thesis further strengthen the call to integrate environmental considerations in economic development and poverty reduction policies. Despite the challenges and uncertainties that lie ahead in an era of climate change, this thesis has identified opportunities to support the livelihoods of the poorest, protect the environment, and build more resilient economic systems.

Appendices

Appendix A (Chapter 1)

Table 1. Observation / Year / Country Table

ISO3 country code	1996	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014
afg									34				34			
arm														11		
aut													9			
bgd		7					7					7				
bgr												28	28			
bih						3			3				3			
blr														6	6	
bol													9			
bra		27										27				
btn														18		
chl		14						16			16		14			
cmr			10						10							
cod							8									
col				22	22	22	22			22	22	22	22	22		
cri												1	1	1	1	
dnk													5			
dom		32	32	32	32	32	32	32	32	32	32	32	32			
ecu									15	15	17	17	17	17	17	
est						15							15			
geo													10	7		
gin									8					8		
gtm								22					22			
hti														10		
hun													20	20		
idn				30	30	30	31	33	33	33	33	33	33	33	33	
ind						31						36	35			
irq									18					18		
jor												12				
kaz															14	
kgz											9		9	9	9	
lao				18						18					17	
lbn							6									
lbr									15							
lva													5	5		
mar			13						13							
mda													33			
mex										32		32		32		
mne													20	20		
moz	11				11					11						
nga						37						37				
nld													12			
npl												5				
pan										9	9	9	12	12	12	
pse					2		2	2			2	2	2			
rou												42	42			
rus				1									83	84		
sdn											15					
sle					3								3			
slv							14	14	14	14	14	14	14	14		
ssd											10					
svk													8			
svn												12	12			
swe													21			
syr									14							
tgo								5					5			
tjk											5					
tls									13							13
tun		21					21					21				
ven				24	24	24	24	24	24	24	24	24	24	24	24	
vnm												65		65		
yem							20									
zmb								9				9				

Table 2. First stage – the effect of rainfall on NPP

	(1)	(2)	(3)	(4)
Variables	IV	IV - Rural	IV - Rural no outliers	IV - Rural SSA
Log Precipitation	0.62***	0.63***	0.64***	1.74***
	(0.04)	(0.07)	(0.08)	(0.19)
Constant	-3.57***	-3.64***	-3.66***	-8.37***
	(0.18)	(0.56)	(0.57)	(0.92)
Observations	2,738	1,362	1,306	104
R-squared	0.81	0.81	0.80	0.64
Country FE	YES	YES	YES	NO
Year FE	YES	YES	YES	NO

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 3. First stage – the effect of precipitation on top soil carbon

	(1)	(2)	(3)	(4)
Variables	IV	IV - Rural	IV - Rural no outliers	IV - Rural SSA
Log Precipitation	0.27***	0.20***	0.25***	0.34***
	(0.03)	(0.04)	(0.05)	(0.08)
Constant	2.03***	2.18***	1.98***	1.90***
	(0.13)	(0.19)	(0.22)	(0.42)
Observations	933	476	452	64
R-squared	0.74	0.74	0.71	0.65
Country FE	YES	YES	YES	NO
Year FE	YES	YES	YES	YES

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 4. Second stage – the effect of top soil carbon on poverty including controls

	(1)	(2)	(3)	(4)
Variables	OLS	IV	IV- Rural	IV - Rural no outliers
Log Top soil carbon	-4.63***	-0.92	-24.43***	-16.57**
	(1.09)	(5.72)	(7.69)	(8.35)
Log Precipitation	0.76			
	(1.24)			
Log Ruggedness	0.37	0.62	-0.87	0.10
	(0.40)	(0.55)	(0.72)	(0.79)
Log Road density	-3.25***	-3.19***	-3.08**	-3.20**
	(0.75)	(0.75)	(1.47)	(1.41)
Share cropland	5.33*	5.76**	-8.83	-6.38
	(3.17)	(2.87)	(6.27)	(6.58)
Share urban	-5.33	-4.37	-36.63**	-33.96**
	(7.56)	(7.18)	(18.02)	(16.81)
Share grassland	1.27	2.23	5.08	6.32
	(4.48)	(4.34)	(10.22)	(10.03)
Share forest	4.41	3.74	-11.13	-11.81
	(3.25)	(3.62)	(7.35)	(7.46)
Log Population	-1.24**	-1.31***	-0.14	-0.06
	(0.49)	(0.47)	(0.72)	(0.71)
Andisol	5.18**	3.19	14.21***	11.31**
	(2.46)	(4.04)	(5.00)	(4.59)
Ardisol	-0.56	-0.23	-0.33	-3.85
	(2.35)	(2.46)	(5.85)	(9.38)
Entisol	0.56	0.47	2.01	2.56
	(1.75)	(1.66)	(2.11)	(2.11)
Gelisol	-5.89*	-7.45**		
	(3.13)	(3.71)		
Histosol	4.66**	0.73	22.68***	16.05*
	(2.31)	(6.57)	(7.77)	(8.23)
Inceptisol	-1.27	-1.97	1.01	0.57
	(1.17)	(1.58)	(1.67)	(1.58)
Mollisol	1.01	0.45	2.88*	2.21
	(1.13)	(1.17)	(1.49)	(1.42)
Oxisol	-6.39**	-6.40**	-0.43	-2.53
	(2.91)	(2.78)	(3.96)	(3.88)
Rock	3.41	2.97	-1.82	-1.56
	(3.19)	(3.06)	(5.30)	(6.20)
Sand	2.96	3.43	-18.77***	-16.20***
	(6.37)	(5.78)	(4.46)	(4.97)
Spodosol	-2.23	-4.12	5.94	3.70
	(1.83)	(3.46)	(4.52)	(3.94)
Ultisol	1.34	1.32	4.36*	3.99*
	(1.90)	(1.83)	(2.47)	(2.41)

Vertisol	2.13	1.44	0.39	-1.19
	(2.75)	(2.91)	(2.88)	(2.83)
Constant	65.17***	53.55**	144.29***	107.28***
	(10.48)	(23.31)	(30.76)	(32.18)
Observations	932	932	475	451
R-squared	0.77	0.76	0.76	0.80
Country FE	YES	YES	YES	YES
Year FE	YES	YES	YES	YES

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 5. Second stage – the effect of top soil carbon on GDP per capita including controls

VARIABLES	(1) OLS	(2) IV	(3) IV - Rural	(4) IV - Rural no outliers
Log Top soil carbon	0.14**	-0.27	0.57*	0.58
	(0.06)	(0.24)	(0.32)	(0.36)
Log Precipitation	-0.10			
	(0.06)			
Log Ruggedness	-0.08***	-0.11***	-0.03	-0.03
	(0.02)	(0.03)	(0.03)	(0.04)
Log Road density	0.08*	0.07*	0.18***	0.17**
	(0.04)	(0.04)	(0.07)	(0.07)
Share cropland	-0.70***	-0.81***	-0.13	-0.13
	(0.16)	(0.16)	(0.23)	(0.29)
Share urban	1.41***	1.20***	2.77***	2.80***
	(0.38)	(0.37)	(0.54)	(0.60)
Share grassland	-0.46*	-0.55**	-0.35	-0.33
	(0.27)	(0.26)	(0.42)	(0.46)
Share forest	-0.32*	-0.27	0.34	0.36
	(0.17)	(0.19)	(0.24)	(0.28)
Log Population	-0.93***	-0.92***	-0.93***	-0.93***
	(0.02)	(0.03)	(0.04)	(0.04)
Andisol	-0.19**	0.02	-0.44**	-0.44**
	(0.09)	(0.15)	(0.18)	(0.19)
Ardisol	-0.10	-0.20	0.54***	0.55***
	(0.14)	(0.17)	(0.19)	(0.19)
Entisol	-0.06	-0.05	-0.09	-0.09
	(0.07)	(0.08)	(0.08)	(0.09)
Gelisol	0.75***	0.91***		
	(0.19)	(0.20)		
Histosol	0.96***	1.37***	0.62**	0.61*
	(0.13)	(0.28)	(0.30)	(0.34)
Inceptisol	0.01	0.09	-0.17**	-0.17**
	(0.06)	(0.08)	(0.07)	(0.07)
Mollisol	0.02	0.09	-0.02	-0.01
	(0.05)	(0.07)	(0.07)	(0.08)
Oxisol	0.37***	0.35***	0.13	0.14
	(0.13)	(0.13)	(0.14)	(0.15)
Rock	-0.24	-0.24	-1.25***	-1.25***
	(0.31)	(0.32)	(0.13)	(0.14)
Sand	-0.22	-0.27	0.07	0.08
	(0.14)	(0.18)	(0.17)	(0.19)
Spodosol	0.20*	0.39**	-0.00	-0.01
	(0.11)	(0.16)	(0.17)	(0.19)
Ultisol	0.09	0.08	0.05	0.06
	(0.08)	(0.08)	(0.08)	(0.08)
Vertisol	0.13	0.13	0.14	0.16

	(0.12)	(0.12)	(0.12)	(0.13)
Constant	10.53***	8.82***	4.06***	4.05**
	(0.51)	(1.08)	(1.41)	(1.62)
Observations	635	635	338	329
R-squared	0.96	0.96	0.98	0.98
Country FE	YES	YES	YES	YES
Year FE	YES	YES	YES	YES

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Appendix B (Chapter 2)

Table 6. Summary statistics of consumption, environmental risk and control variables, 2010, 2012, 2014

Variable	Description	Type	Pooled Cross-Section			Panel Dataset	
			2010	2012	2014	2010-2012	2012-2014
Number of households			9,398	9,399	9,399	4,134	3,963
Consumption	Total per capita expenditure in 2011 PPP	Household - Time variant	2,965	3,188	3,395	258	236
Environmental risks							
Air pollution	Area-weighted mean PM2.5 concentration 2000-10 in micrograms per m3	Commune - Fixed	17.0	17.1	17.0	17.3	17.3
Tree cover loss	Share of forest area affected by tree cover loss in 2001-10	Commune - Fixed	1.5	1.5	1.5	1.6	1.5
Land degradation	Share of land area affected by significant biomass decline	Commune - Fixed	17.9	17.8	17.9	17.2	17.7
Slope	Area-weighted average of slope categories: 1=least steep - 8=most steep	Commune - Fixed	3.37	3.38	3.37	3.41	3.39
Rainfall variability	Standard deviation of monthly rainfall 1981-2010	Commune - Fixed	55.2	54.6	55.8	54.9	56.4
Temperature variability	Standard deviation of mean annual temperature 1981-2010	Commune - Fixed	0.72	0.74	0.74	0.74	0.74
Flood hazard	Share of area at risk of a flood event with a 25 year return period	Commune - Fixed	32.1	32.0	32.0	31.8	31.8
Drought hazard	Number of months to overcome the maximum accumulated deficit volume	Commune - Fixed	1.03	1.03	1.03	1.04	1.03
Control variables							
Current rainfall	Average of monthly rainfall in survey year in mm	Commune- Time variant	142	183	169	41.3	-14.0
Current temperature	Annual average of monthly mean temperature in survey year in Celsius degree	Commune- Time variant	26.0	25.5	25.2	-0.58	-0.13
Long-term rainfall	Mean of monthly rainfall in last 30 years	Commune - Fixed	157	156	158	156	158
Long-term temperature	Mean of monthly temperature mean in last 30 years	Commune - Fixed	25.4	25.4	25.5	25.4	25.4
Distance city	Distance from commune to next main city	Commune - Fixed	29.9	29.9	29.9	30.2	30.0
Distance road	Distance from commune to next road	Commune - Fixed	3.04	3.04	3.03	3.07	3.07
Area agriculture	Area for agricultural activities household has access to	Household - Time variant	3.86	3.99	3.93	0.18	0.09
Area forest	Forest area household has access to	Household - Time variant	1.83	1.47	1.34	0.36	-0.08
Area water surface	Water surface area household has access to	Household - Time variant	0.38	0.34	0.35	-0.02	0.02
Workforce	Share of household members involved in income generating activities	Household - Time variant	0.79	0.78	0.79	0.00	0.01
Education	Average number of school years of household members	Household - Time variant	6.48	6.56	6.66	0.12	0.02
Age head	Age of household head in years	Household - Fixed	48.3	49.7	50.7	50.4	51.3
Female head	Dummy = 1 if household head is female	Household - Fixed	0.25	0.25	0.26	0.25	0.25
Minority	Dummy = 1 if household belongs to ethnic minority	Household - Fixed	0.17	0.17	0.17	0.18	0.17

Notes: The table indicates mean values for each year for the Pooled Cross-Section. For the Panel Dataset, mean differences between years are presented for time variant variables and mean values for fixed variables.

Table 7. Number of poor people in 2010 across environmental risk categories at district level

Risk category	Air pollution	Tree cover loss	Land degradation	Slope	Rainfall Variability	Temperature Variability	Flood hazards	Drought hazards
Low	23,393	19,792	18,206	21,302	22,117	21,688	30,318	20,172
Medium	24,874	25,046	27,650	21,858	28,054	27,455	21,769	24,281
High	25,741	28,967	28,713	30,632	24,106	25,124	21,704	29,349
ANOVA								
F	1.08	16.99	28.63	22.27	7.09	6.56	19.82	16.94
Prob>F	0.3405	0.0000	0.0000	0.0000	0.0009	0.0015	0.0000	0.0000

Notes: The table shows the number of poor people across the three environmental risks categories as calculated for Figure 4 and statistics from a one-way analysis-of-variance (ANOVA), which assess whether the difference in poverty rates across risk categories is statistically significant.

Table 8. Environmental risks across socio-economic groups in 2014

	Air pollution	Treecover loss	Land degradation	Slope	Rainfall variability	Temperature variability	Flood hazards	Drought hazards
Non-minority	16.41	1.2	17.58	2.89	56.32	0.72	35.97	1.02
Minority	20.1	3.12	19.5	5.81	53.43	0.84	12.11	1.11
Pr(|T| > |t|)	0.00	0.00	0.04	0.00	0.00	0.00	0.00	0.00
Non-poor (88%)	16.76	1.31	18.07	3.13	55.58	0.73	33.88	1.03
Poor (12%)	19.17	3.15	16.50	5.27	58.00	0.82	17.08	1.06
Pr(|T| > |t|)	0.00	0.00	0.15	0.00	0.00	0.00	0.00	0.05
T60 (60%)	16.34	1.20	18.28	3.01	54.82	0.72	33.33	1.01
B40 (40%)	18.02	1.98	17.35	3.89	57.36	0.77	30.16	1.07
Pr(|T| > |t|)	0.00	0.00	0.18	0.00	0.00	0.00	0.00	0.00
Urban	15.73	0.79	20.86	3.08	53.86	0.72	24.50	0.97
Rural	17.56	1.82	16.66	3.49	56.67	0.75	35.20	1.06
Pr(|T| > |t|)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Notes: The table indicates the mean value for the group and probability from two-sample t-test for difference between the two groups.

Table 9. Correlation coefficients between consumption in 2014 and environmental risk variables

	Air pollution	Treecover loss	Land degradation	Slope	Rainfall variability	Temperature variability	Flood hazards	Drought hazards
All Households	-0.04	-0.09	0.01	-0.15	-0.03	-0.07	-0.03	-0.09
Significance level	0.00	0.00	0.41	0.00	0.00	0.00	0.01	0.00
Rural Households	-0.07	-0.08	0.01	-0.22	-0.03	-0.14	0.12	-0.05
Significance level	0.00	0.00	0.53	0.00	0.01	0.00	0.00	0.00
Urban Households	0.06	-0.05	-0.03	-0.04	0.01	0.05	-0.10	-0.08
Significance level	0.00	0.01	0.14	0.04	0.68	0.01	0.00	0.00

Notes: The table shows the correlation coefficient between per-capita consumption and the environmental risk variables for all households and the rural and urban sub-sample.

Table 10. The effect of environmental risks on consumption differences between 'Pooled' households in 2010, 2012 and 2014

Dependent variable: Ln of per-capita expenditure							
	All	Poor	B40	Rural	Rural 2014	Urban	Urban 2014
Air pollution	0.000511	0.00176	**0.00255****	**0.00425*****	0.00220	-0.00178	-0.00315
	(0.32)	(1.01)	(2.18)	(2.60)	(1.14)	(-0.55)	(-0.81)
Tree cover loss	-0.00179	0.00229	-0.00125	0.00132	-0.00248	-0.00705	-0.00500
	(-1.18)	(1.60)	(-1.10)	(0.88)	(-1.37)	(-1.35)	(-0.74)
Land degradation	**0.000300***	-0.000256	0.000171	0.000108	0.000197	-0.0000844	-0.00000728
	(1.87)	(-1.27)	(1.26)	(0.55)	(0.81)	(-0.33)	(-0.02)
Slope	**-0.0213*****	**-0.0148****	**-0.0102****	**-0.0101***	-0.00389	-0.0164	-0.0117
	(-3.98)	(-2.32)	(-2.43)	(-1.71)	(-0.55)	(-1.58)	(-0.88)
Rainfall variability	**-0.00276*****	-0.00120	-0.000420	**-0.00281*****	**-0.00499*****	-0.000672	-0.000577
	(-4.48)	(-1.46)	(-0.75)	(-4.21)	(-4.73)	(-0.58)	(-0.33)
Temperature variability	**-0.265***	0.0382	**-0.212***	**-0.454*****	-0.0311	0.113	0.318
	(-1.88)	(0.22)	(-1.68)	(-2.98)	(-0.17)	(0.42)	(0.95)
Flood hazard	**-0.00144*****	-0.000202	0.00000600	0.000183	0.000441	**-0.00184*****	**-0.00201*****
	(-5.97)	(-0.74)	(0.04)	(0.67)	(1.43)	(-4.05)	(-3.96)
Drought hazard	**-0.0335*****	0.00353	**-0.0199****	**-0.0428*****	**-0.0379****	-0.0380	-0.0477
	(-2.67)	(0.26)	(-2.01)	(-3.17)	(-2.29)	(-1.59)	(-1.56)
<u>Control variables</u>							
Current rainfall	0.00121***	0.000473*	0.000247	0.00134***	0.00202***	0.0000388	0.000273
	(5.32)	(1.68)	(1.21)	(5.36)	(4.35)	(0.09)	(0.30)
Current temp	0.0322	-0.0504**	-0.0400**	0.00195	0.00123	0.0381	0.120
	(1.55)	(-2.20)	(-2.10)	(0.09)	(0.03)	(0.97)	(1.45)
Long-term rain	-0.00191**	-0.000913	-0.00109*	-0.00181**	-0.000495	-0.000980	-0.00148
	(-2.37)	(-0.96)	(-1.66)	(-2.23)	(-0.49)	(-0.60)	(-0.68)
Long-term temp	-0.0380*	0.0658***	0.0586***	0.00675	0.0270	-0.0321	-0.0998
	(-1.70)	(2.67)	(2.87)	(0.28)	(0.59)	(-0.78)	(-1.18)
Distance city	-0.00230***	-0.000223	-0.000399	-0.00101***	-0.000591	-0.00253***	-0.00220***
	(-7.79)	(-0.76)	(-1.63)	(-3.32)	(-1.51)	(-4.79)	(-3.32)
Distance road	-0.0110***	-0.00267**	-0.00462***	-0.00832***	-0.00785***	0.00329	0.00618
	(-6.29)	(-2.13)	(-4.01)	(-5.57)	(-4.12)	(0.62)	(1.05)
Area agriculture	0.00153***	0.00130*	0.00296***	0.00426***	0.00396***	-0.00169	-0.000953
	(2.93)	(1.67)	(5.52)	(7.95)	(5.68)	(-1.24)	(-0.61)
Area forest	-0.00000525	-0.000144**	-0.0000523	0.0000935	-0.000213	0.0000408	0.00505
	(-0.05)	(-2.46)	(-0.34)	(1.43)	(-0.49)	(0.01)	(0.88)
Area water	0.00467***	0.00170	0.00272***	0.00432***	0.00753***	0.00517**	0.00691
	(3.31)	(1.02)	(2.81)	(3.21)	(5.16)	(2.15)	(1.03)
Workforce	-0.115***	0.0910***	0.0735***	-0.00266	0.00383	-0.168***	-0.151***
	(-10.90)	(6.69)	(8.39)	(-0.24)	(0.25)	(-7.41)	(-4.56)
Education	0.103***	0.0243***	0.0375***	0.0860***	0.0849***	0.108***	0.105***
	(63.51)	(10.75)	(23.50)	(49.28)	(33.60)	(36.69)	(23.40)
Age head	0.000656**	0.000413	0.000315	0.000497	-0.0000227	0.000576	0.00140*
	(2.43)	(1.34)	(1.34)	(1.64)	(-0.05)	(1.12)	(1.91)
Female head	0.0854***	-0.00897	0.00950	0.0684***	0.0673***	0.0399***	0.0303
	(9.56)	(-0.69)	(1.11)	(6.36)	(4.37)	(2.74)	(1.47)
Minority	-0.409***	-0.0948***	-0.228***	-0.401***	-0.429***	-0.290***	-0.346***

	(-24.35)	(-6.11)	(-17.64)	(-22.16)	(-17.71)	(-7.87)	(-6.69)
Year 2012	0.0607***	-0.000678	0.0835***	0.0558***		0.0660**	
	(3.59)	(-0.03)	(5.33)	(3.02)		(2.13)	
Year 2014	0.150***	-0.0142	0.155***	0.147***		0.113***	
	(7.68)	(-0.57)	(8.51)	(6.99)		(3.13)	
_cons	8.024***	6.487***	6.849***	7.513***	6.677***	7.580***	7.225***
	(35.15)	(25.99)	(32.66)	(30.26)	(20.11)	(18.23)	(13.23)
N	27698	3811	11025	19748	6492	7950	2715
R-sq	0.432	0.204	0.365	0.399	0.409	0.316	0.305

Notes: The table indicates coefficients estimated from 'Pooled' cross-section model using Ordinary Least Squares. *
0.10 ** 0.05 *** 0.01 significance level. Values in parentheses indicate standard errors corrected for cluster correlation at commune-level. B40 denotes households in the bottom two consumption quintiles.

Table 11. The effect of environmental risks on consumption changes over time of 'Panel' households in 2010-12 and 2012-14

Dependent variable: Change in per-capita expenditure							
	All	Poor	B40	Rural	Rural 2012-14	Urban	Urban 2012-14
Air pollution	**-12.37***	6.040	-0.0273	-4.189	-3.224	-34.10	-41.43
	(-1.91)	(1.52)	(-0.01)	(-0.86)	(-0.42)	(-1.58)	(-1.61)
Tree cover loss	-4.166	-1.376	-5.865	-4.605	**-13.16***	-0.818	-5.808
	(-0.73)	(-0.40)	(-1.63)	(-0.83)	(-1.92)	(-0.03)	(-0.11)
Land degradation	0.262	0.554	0.0442	**1.555****	0.542	-2.124	-2.237
	(0.33)	(0.96)	(0.09)	(2.13)	(0.47)	(-1.15)	(-0.72)
Slope	8.262	-17.99	4.052	-2.692	-1.030	38.90	67.10
	(0.43)	(-0.93)	(0.30)	(-0.16)	(-0.04)	(0.73)	(0.91)
Rainfall variability	-0.162	**-3.139***	0.445	-0.349	-0.132	-0.250	**-15.18***
	(-0.08)	(-1.71)	(0.32)	(-0.17)	(-0.04)	(-0.04)	(-1.72)
Temperature variability	**974.5***	193.5	182.0	377.8	403.0	2582.1	**6247.0****
	(1.74)	(0.61)	(0.67)	(0.88)	(0.63)	(1.42)	(2.40)
Flood hazard	-0.390	0.110	**1.533****	-0.0430	-0.602	-4.244	-3.479
	(-0.38)	(0.10)	(2.33)	(-0.05)	(-0.45)	(-1.30)	(-0.99)
Drought hazard	-2.442	**-121.1****	-64.84	-35.65	53.63	137.7	58.11
	(-0.05)	(-2.06)	(-1.55)	(-0.72)	(0.67)	(0.89)	(0.23)
Control variables							
Current rainfall	1.259	0.656	-0.458	1.625	0.635	-0.560	2.514
	(1.09)	(0.89)	(-0.66)	(1.59)	(0.40)	(-0.16)	(0.50)
Current temperature	-190.0**	-37.60	-48.73	-103.5	-40.98	-406.5*	-239.6
	(-2.20)	(-0.57)	(-0.90)	(-1.32)	(-0.22)	(-1.69)	(-0.46)
Long-term rainfall	1.080	5.660***	1.731	-0.469	0.421	6.838	24.23*
	(0.41)	(2.69)	(1.03)	(-0.20)	(0.10)	(0.85)	(1.89)
Long-term temperature	54.08**	-44.43***	9.712	25.99	59.33*	109.1**	226.8***
	(2.48)	(-2.74)	(0.74)	(1.20)	(1.93)	(2.04)	(2.85)
Distance city	2.372**	0.965	1.647**	1.450	1.915	3.100	0.678
	(2.24)	(1.33)	(2.27)	(1.61)	(1.28)	(1.20)	(0.17)
Distance road	5.620	3.284	5.966**	2.183	17.50***	15.10	13.03
	(1.15)	(1.13)	(1.97)	(0.49)	(2.64)	(0.65)	(0.47)
Area agriculture	13.91***	8.732**	14.39	12.56***	8.157**	19.83	39.26*
	(2.59)	(2.17)	(1.64)	(2.62)	(2.24)	(1.09)	(1.93)
Area forest	0.336	0.146	0.155	0.315	0.209	22.92	35.49
	(0.28)	(0.09)	(0.17)	(0.28)	(0.13)	(0.63)	(0.72)
Area water surface	25.81**	24.64	11.20	33.36***	36.21***	-54.07	-130.1
	(2.01)	(0.89)	(1.47)	(3.80)	(4.73)	(-1.18)	(-1.53)
Workforce	190.6***	166.7***	89.93**	208.4***	228.1***	127.9	-489.4
	(2.63)	(2.82)	(2.17)	(3.42)	(2.72)	(0.44)	(-1.08)
Education	117.4***	40.56	67.02***	73.88***	88.50***	235.3***	203.1***
	(5.75)	(1.40)	(5.05)	(4.26)	(3.77)	(4.02)	(3.35)
Age head	-2.285	-3.548***	-2.276**	-1.553	-4.078*	-4.467	-5.595
	(-1.60)	(-3.09)	(-2.53)	(-1.14)	(-1.94)	(-1.18)	(-0.95)
Female head	85.89	28.10	30.55	49.79	26.04	193.3	385.3*
	(1.52)	(0.64)	(0.91)	(1.03)	(0.33)	(1.44)	(1.94)
Minority	-152.8***	175.7***	234.9***	-149.6**	-204.7**	-179.1	-702.7**
	(-2.63)	(3.27)	(5.47)	(-2.56)	(-2.09)	(-0.80)	(-2.09)
Year 2012-14	139.8	-101.7	-128.8**	96.08		202.4	
	(1.35)	(-1.46)	(-2.24)	(1.04)		(0.71)	
_cons	-1943.4**	188.1	-784.5*	-542.3	-1461.6	-5280.7**	-12451.3***
	(-2.25)	(0.39)	(-1.70)	(-0.73)	(-1.38)	(-2.17)	(-3.05)
N	7957	952	3250	5804	2792	2153	1109
R-sq	0.017	0.171	0.076	0.020	0.035	0.029	0.050

Notes: The table indicates coefficients estimated from 'Panel' model using Ordinary Least Squares and differences over time for time-variant variables. * 0.10 ** 0.05 *** 0.01 significance level. Values in parentheses indicate standard errors corrected for cluster correlation at commune-level. B40 denotes households in the bottom two consumption quintiles.

Appendix C (Chapter 3)

Appendix C1

National-level: Flood Hazard Modeling Details

All the data were produced using the SSBN global flood model, producing flood hazard data at 90m resolution. The SSBN global model couples a flood frequency analysis conducted at the global scale, with a fully 2-D hydraulic modeling framework. Extreme river discharges are derived from a Flood Frequency Analysis (FFA), applied at the global scale (Smith, Sampson and Bates, 2015). The model also explicitly simulates in-channel flow, with the FFA also used to estimate bankfull discharge across the channel network (defined as the 1.1 – 2 year event depending on climate zone); these values are used to calibrate the channel conveyance capacity within the hydraulic modeling framework. A number of global data sets are used to derive the inputs to the hydraulic model. Firstly, the Hydrosheds variant of SRTM is used, both at 3 and 30 arc second resolutions. A number of additional corrections are applied to the terrain data including a systematic vegetation correction procedure in vegetated areas and an urban correction procedure in urbanized areas. A detailed description of the modeling framework is provided by (Sampson *et al.*, 2015).

For the coastal simulations, input boundary conditions were derived using estimates of return period surge heights, taken from (Van Chinh, JianCheng and Trinh, 2014). Storm surge hydrographs for each recurrence interval were taken for four tidal gauges located along the Vietnamese coastline. Coastal boundary conditions for the hydraulic model were derived by linearly interpolating between the gauge locations. The hydraulic model was set up so that a coastal boundary condition was implemented for each 'land' cell located next to the coast. In addition to the coastal boundary conditions, large river channels were also included in the simulations, using a sub-grid channel network set-up (Neal, Schumann and Bates, 2012). In channel flow was estimated to be 0.5*bankfull Q; rivers were estimated to be at 50 percent channel capacity.

Coastal simulations under future climate conditions were undertaken using the latest projections of global mean sea-level rise, outlined in the Fourth Assessment Report (AR5) and Fifth Assessment Report (AR5) of the Intergovernmental Panel on Climate Change (IPCC) (IPCC, 2007, 2014). Estimates of sea-level changes were taken and used to directly perturb the boundary conditions used in the simulations under current conditions. In order to incorporate uncertainty, simulations were undertaken for a range of projected changes, represented here as Low, Medium and High sea-level rise (SLR) projections (Table 15 in the text). The simulations are all conducted assuming that no flood defenses are in place; clearly flood defenses are not represented in the available terrain data. Therefore, these simulations should be considered as an upper bound of flood exposure in the country.

The possibility of including a storm surge intensity component to the future projections was also discussed, but there are significant uncertainties around quantifying storm intensity changes that would preclude any reasonable modeling being undertaken (Seneviratne *et al.*, 2012). It also seems that changes in surge extremes are going to be largely driven by sea-level rise (Lowe *et al.*, 2010). That being said, there are studies attempting to quantify changes in storm surge intensity; Lin *et al.* (2012) reported in a study focused on the North Atlantic, that in some cases changing storm surge intensity was comparable to changing sea-level rise. Such changes would effectively double change in hazard intensities presented here. As of yet, we have not included this in the simulations due to the uncertainty.

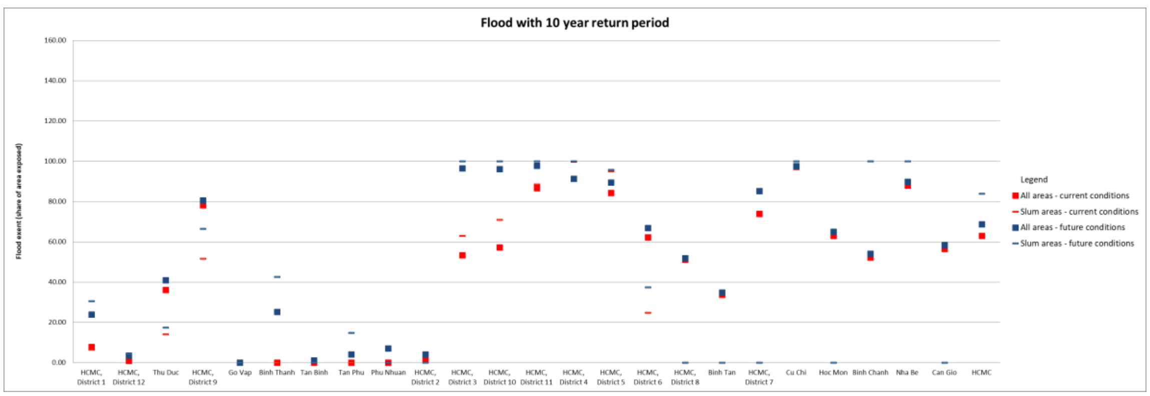
Appendix C2

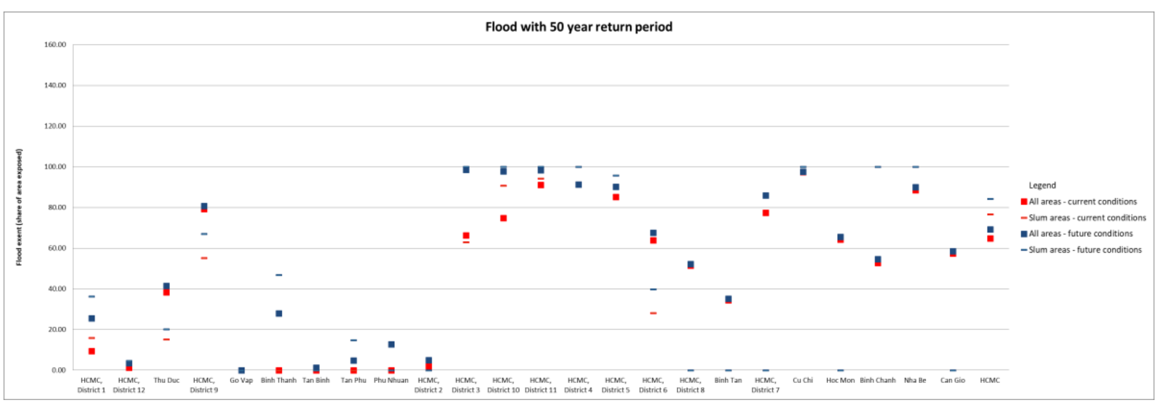
City-case: Ho Chi Minh City

Results for flood extent across districts

Figure 1 presents the differences in relative flooding extent (percent) for a flooding event with a 10 year or 50 year return period representing the current and future conditions. Overall, we find that both under current conditions and given a 30 cm sea level rise, relative flooding extent is higher in slum areas than in the non-slum areas, both when looking at the HCMC-city level and at the level of districts. This can serve as a first-order indicator that these slum areas are relatively often located in flood prone areas.

Figure 1. Graph showing the differences in relative flooding extent (from 0 to 100% exposure) for a (a) 1/10 and (b) 1/50 year flooding event under current and future conditions per district and for the whole city of HCMC between all areas and urban slum areas.

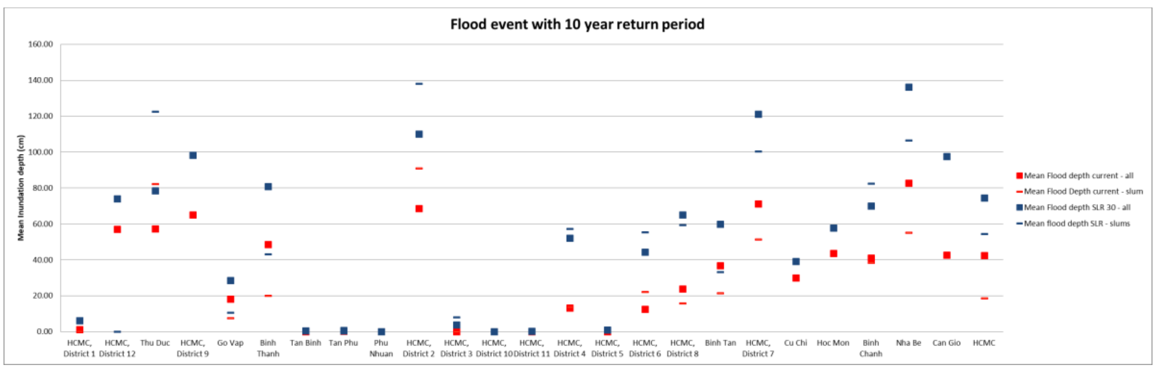


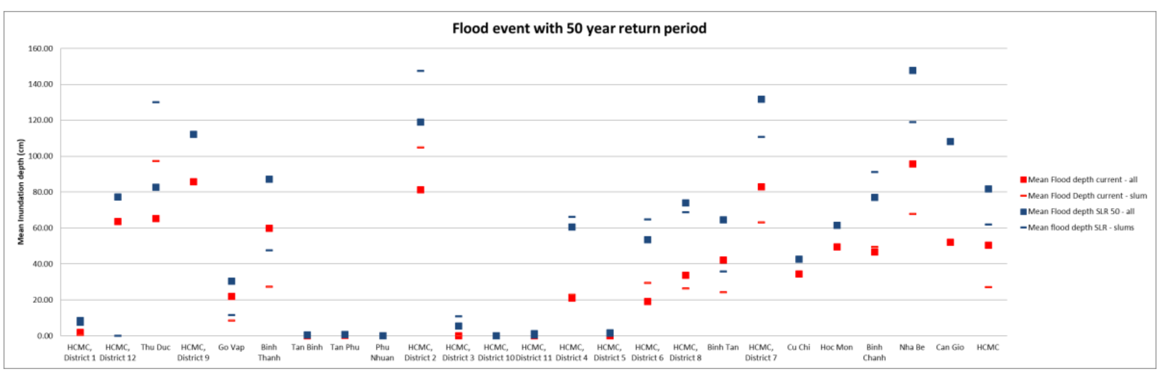


Flood depth analysis

We also examine the exposure to flooding in terms of mean flooding depth. At the city-level, mean inundation depths were found to be higher in the urban non-slum area compared to the slum locations under any of the return periods used (Figure 2). However, spread (standard deviation) in inundation depths was found to be very large when looking at the city-totals. Looking at the individual districts, we find a higher inundation depth within slums – compared to non-slum areas – in five districts for the 10-year return period flood and up to eight districts for the 1000-year return period flood. A sea-level rise of 30cm increased mean inundation depths for the entire city by 30–40 cm depending on the return period. No significant differences were found in the increase in inundation depths between slum and non-slum areas.

Figure 2. Graph showing the difference in mean inundation depth (from 0 to 180cm) between slum areas, and all areas, per district and for the whole city of HCMC using a flooding scenario with a (a) 1/10 and (b) 1/50 year return period representing current and future conditions.

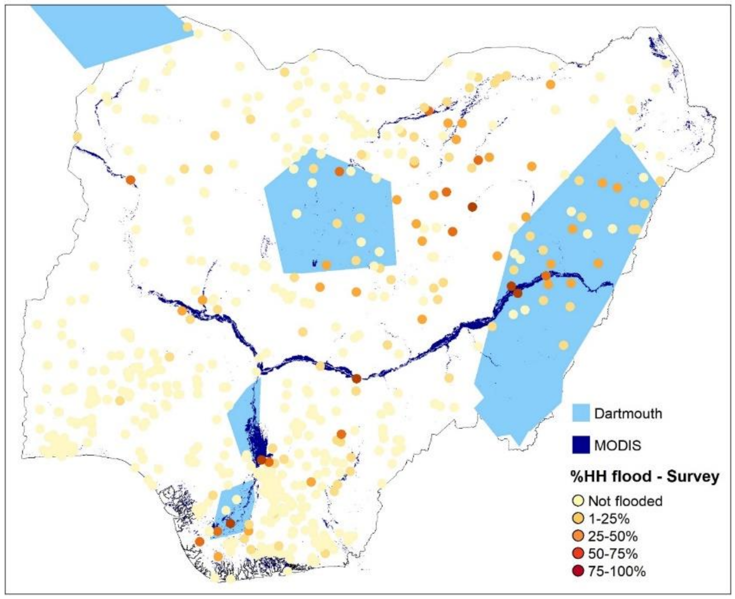




Appendix D (Chapter 4)

Appendix D1

Nigeria's flood extent from 2012 delineated from satellite imagery from multiple sources



Notes: Sources: (Nzeribe, Nwokoye and Ezenekwe, 2014; Dartmouth Flood Observatory, 2021; NASA, 2023)

Appendix D2

Main results with inclusion of controls

Impact of flood on agricultural production, for all production and the major crops.

	(1) All Crops	(2) Maize	(3) Sorghum	(4) Cassava	(5) Bean	(6) Yams	(7) Mill	(8) Rice
Flood	--689.4*	3.1	-438.5***	-418.0**	-26.8	869.8	-168.5*	-83.8
	(362.7)	(106.2)	(105.0)	(199.2)	(41.1)	(568.4)	(96.0)	(71.4)
Obs	768	376	416	182	358	162	290	144
N of HH	384	188	208	91	179	81	145	72
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
HH FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Cluster	HHID	HHID	HHID	HHID	HHID	HHID	HHID	HHID
Outliers	W95	W95	W95	W95	W95	W95	W95	W95

Note: Controls include number of household members, bank account presence, landholdings per capita, distance to market, and slope of plot. Results include household fixed effects and time fixed effects and are clustered at the household level. W95 indicates the dependant variable is top-end windsorized at the 95th percentile. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Impacts of the flood on crop value for affected farmers.

	(1) Crop Value
Flood	-69,996***
	(12,756)
Obs	768
N of HH	384
Controls	Yes
HH FE	Yes
Time FE	Yes
Cluster	HHID
Outliers	W95

Notes: Controls include number of household members, bank account presence, landholdings per capita, distance to market, and slope of plot. Crop value metric is measured in Naira. Results include household fixed effects and time fixed effects and are clustered at the household level. W95 indicates the dependant variable is top-end windsorized at the 95th percentile. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Effect of living in a flooded area on crop value, for households who also report being individually affected (Column 2) versus those who do not report being individually affected (Column 3).

	(1) Yes individually affected	(2) No individually affected
Flood	-39,300***	31,187***
(Satellite)	(13,786)	(10,757)
Obs	768	768
N of HH	384	384
Controls	Yes	Yes
HH FE	Yes	Yes
Time FE	No	No
Cluster	HHID	HHID
Outliers	W95	W95

Note: Controls include number of household members, bank account presence, landholdings per capita, distance to market, and slope of plot. Crop value metric is measured in Naira. Results include household fixed effects and time fixed effects and are clustered at the household level. W95 indicates the dependant variable is top-end windsorized at the 95th percentile. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Time fixed effects removed due to collinearity.

Impact of the flood on total consumption per capita for all households.

VARIABLES	(1) Total Consumption	(2) Food Consumption
Flood	-1,753	-1,678
	(6,008)	(4,428)
Observations	768	768
Number of hhid	384	384
Controls	Yes	Yes
Household FE	Yes	Yes
Time FE	Yes	Yes
Cluster	HHID	HHID

Notes: Controls include number of household members, bank account presence, landholdings per capita, distance to market, and slope of plot. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Impact of flood on livestock value and livestock sales for agricultural households.

VARIABLES	(1) Livestock value	(2) Value of sales
Flood	-22,445	-1,341
	(28,109)	(1,594)

Observations	768	768
Number of hhid	384	384
Controls	Yes	Yes
Household FE	Yes	Yes
Time FE	Yes	Yes
Cluster	HHID	HHID

Note: Controls include number of household members, bank account presence, landholdings per capita, distance to market, and slope of plot. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Livestock value and value of sales are represented in Naira.

Appendix D3

Impact of flood on agricultural production, for all production and the major crops.

	(1) All Crops	(2) Maize	(3) Sorghum	(4) Cassava	(5) Bean	(6) Yams	(7) Mill	(8) Rice
Flood	-250.5	-176.6	-178.6	-106.6	-31.2	5093.8***	-136.3	-229.9*
	(620.6)	(227.3)	(194.3)	(319.8)	(58.3)	(985.8)	(150.0)	(115.0)
Obs	364	160	176	106	166	84	150	48
N of HH	182	80	88	53	83	42	75	24
Controls	No	No	No	No	No	No	No	No
HH FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Cluster	HHID	HHID	HHID	HHID	HHID	HHID	HHID	HHID
Outliers	W95	W95	W95	W95	W95	W95	W95	W95

Note: Controls include number of household members, age of household head, bank account presence, landholdings per capita, distance to market, slope of plot, and elevation of plot. Results include household fixed effects and time fixed effects and are clustered at the household level. W95 indicates the dependant variable is top-end windsorized at the 95th percentile. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Impacts of the flood on crop value for affected farmers.

	(1) Crop Value
Flood	-40,019**
	(17,380)
Obs	364
N of HH	182
Controls	No
HH FE	Yes
Time FE	Yes
Cluster	HHID
Outliers	W95

Notes: Controls include number of household members, age of household head, bank account presence, landholdings per capita, distance to market, slope of plot, and elevation of plot. Crop value metric is measured in Naira. Results include household fixed effects and time fixed effects and are clustered at the household level. W95 indicates the dependant variable is top-end windsorized at the 95th percentile. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Effect of living in a flooded area on crop value, for households who also report being individually affected (Column 2) versus those who do not report being individually affected (Column 3).

	(1)	(2)

	Yes individually affected	No individually affected
Flood	-5,224	28,203***
(Satellite)	(14,921)	(10,321)
Obs	364	364
N of HH	182	182
Controls	No	No
HH FE	Yes	Yes
Time FE	No	No
Cluster	HHID	HHID
Outliers	W95	W95

Note: Crop value metric is measured in Naira. Results include household fixed effects and time fixed effects and are clustered at the household level. W95 indicates the dependant variable is top-end windsorized at the 95th percentile. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Time fixed effects removed due to collinearity.

Impact of the flood on total consumption per capita for all households.

VARIABLES	(1) Total Consumption	(2) Food Consumption
Flood	1,809	3,603
	(8,004)	(6,380)
Observations	364	364
Number of hhid	182	182
Controls	No	No
Household FE	Yes	Yes
Time FE	Yes	Yes
Cluster	HHID	HHID

Notes: Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Impact of flood on livestock value and livestock sales for agricultural households.

VARIABLES	(1) Livestock value	(2) Value of sales
Flood	-43,695	-2,978
	(46,067)	(2,576)
Observations	364	364
Number of hhid	182	182
Controls	No	No

Household FE	Yes	Yes
Time FE	Yes	Yes
Cluster	HHID	HHID

Note: Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Livestock value and value of sales are represented in Naira.

Appendix E (Chapter 5)

Appendix E1

Main results when using larger sample

Impact of the conversion program on LPG use

	(1)
	LPG use
Conversion Program	.561***
	(.031)
Observations	1,096
Number of HH	548
Household FE	YES
Year FE	YES

Notes: Households in Lombok (298) are the treated group and households in Sumbawa (250) are the control group. Robust standard errors are included. * p < 0.10, ** p < 0.05, *** p < 0.01.

Impacts of the Conversion Program on fuel and education expenditures

	(1)	(2)
	Fuel	Education
Conversion Program	-10,906***	-46,618***
	(2,865)	(16,246)
Observations	1,096	1,036
Number of HH	548	539
Household FE	YES	YES
Year FE	YES	YES
Outliers	W95	W95

Notes: Households in Lombok (298) are the treated group and households in Sumbawa (250) are the control group. Robust standard errors are included. * p < 0.10, ** p < 0.05, *** p < 0.01. Expenditure figures are in Indonesian Rupiah.

Impacts on the share of women and children under 15 within each household reporting any illness.

	(1)	(2)
Regressor	Women	Children
Conversion Program	-0.026	0.053
	(0.052)	(0.068)

Observations	1,033	752
Number of HH	534	437
Household FE	YES	YES
Year FE	YES	YES

Notes: Figures represent the share of women and children (under 15) within each household reporting any illness in the 4 weeks prior to the survey. Robust standard errors are included. * p < 0.10, ** p < 0.05, *** p < 0.01*

Impacts on the share of women and children under 15 within each household reporting headache illness.

	(1)	(2)
Regressor	Women	Children
Conversion Program	0.024	0.018
	(0.056)	(0.065)
Observations	1,033	752
Number of HH	534	437
Household FE	YES	YES
Year FE	YES	YES

Notes: Figures represent the share of women and children (under 15) within each household reporting headache illness in the 4 weeks prior to the survey. Robust standard errors are included. * p < 0.10, ** p < 0.05, *** p < 0.01*

Impacts on the share of women and children under 15 within each household reporting cough illness.

	(1)	(2)
Regressor	Women	Children
Conversion Program	-0.029	0.001
	(0.051)	(0.065)
Observations	1,033	752
Number of HH	534	437
Household FE	YES	YES
Year FE	YES	YES

Notes: Figures represent the share of women and children (under 15) within each household reporting cough illness in the 4 weeks prior to the survey. Robust standard errors are included. * p < 0.10, ** p < 0.05, *** p < 0.01*

Impacts on the share of women and children under 15 within each household reporting breathing illness.

Regressor	(1) Women	(2) Children
Conversion Program	-0.029	-0.022
	(0.027)	(0.023)
Observations	1,033	752
Number of HH	534	437
Household FE	YES	YES
Year FE	YES	YES

Notes: Figures represent the share of women and children (under 15) within each household reporting breathing illness in the 4 weeks prior to the survey. Robust standard errors are included. * p < 0.10, ** p < 0.05, *** p < 0.01*

Impacts on age-corrected lung capacity for women and children under 15 within each household.

Regressor	(1) Women	(2) Children
Conversion Program	1.96	7.35
	(6.86)	(19.43)
Observations	970	385
Number of HH	515	320
Household FE	YES	YES
Year FE	YES	YES

Notes: Figures represent age-corrected lung capacity for women and the average for children in each household. Lung capacity is measured in mg/L and is collected by a trained surveyor. Robust standard errors are included. * p < 0.10, ** p < 0.05, *** p < 0.01*

Appendix E2

Health outcomes for men and women

All illnesses

Regressor	Any illness - Women	Any illness - Men
LPG Stove Use	-0.078	-0.033
	(0.057)	(0.064)
Observations	804	803
Number of HH	402	402
Household FE	YES	YES
Year FE	YES	YES

Notes: Figures represent the share of women and children (under 15) within each household reporting any illness in the 4 weeks prior to the survey. Robust standard errors are included. * p < 0.10, ** p < 0.05, *** p < 0.01*

Headache

Regressor	Headache - Women	Headache - Men
LPG Stove Use	-0.046	-0.090
	(0.063)	(0.664)
Observations	804	803
Number of HH	402	402
Household FE	YES	YES
Year FE	YES	YES

Notes: Figures represent the share of women and children (under 15) within each household reporting headache illness in the 4 weeks prior to the survey. Robust standard errors are included. * p < 0.10, ** p < 0.05, *** p < 0.01*

Cough

Regressor	Cough illness - Women	Cough illness - Men
LPG Stove Use	-0.119*	-0.104*
	(0.063)	(.059)
Observations	804	803
Number of HH	402	402
Household FE	YES	YES
Year FE	YES	YES

Notes: Figures represent the share of women and children (under 15) within each household reporting cough illness in the 4 weeks prior to the survey. Robust standard errors are included. * p < 0.10, ** p < 0.05, *** p < 0.01*

Breathing

Regressor	Breathing illness - Women	Breathing illness - Men
LPG Stove Use	-0.071**	-0.065*
	(0.036)	(0.038)
Observations	804	803
Number of HH	402	402
Household FE	YES	YES
Year FE	YES	YES

Notes: Figures represent the share of women and children (under 15) within each household reporting breathing illness in the 4 weeks prior to the survey. Robust standard errors are included. * p < 0.10, ** p < 0.05, *** p < 0.01*

Lung capacity

Regressor	Lung capacity - Women	Lung capacity - Men
LPG Stove Use	-3.64	-12.90
	(7.16)	(10.92)
Observations	804	681
Number of HH	402	376
Household FE	YES	YES
Year FE	YES	YES

Notes: Figures represent age-corrected lung capacity for women and the average for children in each household. Lung capacity is measured in mg/L and is collected by a trained surveyor. Robust standard errors are included. * p < 0.10, ** p < 0.05, *** p < 0.01*

References

Adger, W. N. (1999) 'Social vulnerability to climate change and extremes in coastal Vietnam', *World development*, 27(2), pp. 249–269.

Adjei-Mantey, K. and Adusah-Poku, F. (2019) 'Natural disasters and economic growth in Africa'.

Adjei-Mantey, K. and Takeuchi, K. (2021) 'The effect of in utero exposure to household air pollution on child health: evidence from Ghana', *Health Policy OPEN*, 2, p. 100029.

Adjei-Mantey, K., Takeuchi, K. and Quartey, P. (2021) 'Impact of LPG promotion program in Ghana: The role of distance to refill', *Energy Policy*, 158, p. 112578.

Agada, S. and Nirupama, N. (2015) 'A serious flooding event in Nigeria in 2012 with specific focus on Benue State: a brief review', *Natural Hazards*, 77(2), pp. 1405–1414.

Akter, S. and Basher, S. A. (2014) 'The impacts of food price and income shocks on household food security and economic well-being: Evidence from rural Bangladesh', *Global Environmental Change*, 25, pp. 150–162.

Akter, S. and Mallick, B. (2013) 'The poverty–vulnerability–resilience nexus: Evidence from Bangladesh', *Ecological Economics*, 96, pp. 114–124.

Aldrich, D. P. (2012) 'Building resilience', *Social capital in post-disaster recovery*, p. 2012.

Alene, A. D. *et al.* (2018) 'Identifying crop research priorities based on potential economic and poverty reduction impacts: The case of cassava in Africa, Asia, and Latin America', *PLoS One*, 13(8), p. e0201803.

Alexander, D. *et al.* (2017) 'Randomized controlled ethanol cookstove intervention and blood pressure in pregnant Nigerian women', *American journal of respiratory and critical care medicine*, 195(12), pp. 1629–1639.

Alix-Garcia, J. M., Sims, K. R. E. and Yañez-Pagans, P. (2015) 'Only one tree from each seed? Environmental effectiveness and poverty alleviation in Mexico's payments for ecosystem services program', *American Economic Journal: Economic Policy*, 7(4), pp. 1–40.

Amaral, S. *et al.* (2005) 'Estimating population and energy consumption in Brazilian Amazonia using DMSP night-time satellite data', *Computers, Environment and Urban Systems*, 29(2), pp. 179–195.

Andadari, R. K., Mulder, P. and Rietveld, P. (2014) 'Energy poverty reduction by fuel switching. Impact evaluation of the LPG conversion program in Indonesia', *Energy Policy*, 66, pp. 436–449.

Ang, J. B. and Gupta, S. K. (2018) 'Agricultural yield and conflict', *Journal of Environmental Economics and Management*, 92, pp. 397–417.

Angelsen, A. *et al.* (2014) 'Environmental Income and Rural Livelihoods: A Global-Comparative Analysis', *World Development*, xx. doi: 10.1016/j.worlddev.2014.03.006.

Angima, S. D. *et al.* (2003) 'Soil erosion prediction using RUSLE for central Kenyan highland conditions', *Agriculture, ecosystems & environment*, 97(1–3), pp. 295–308.

Angrist, J. D. and Pischke, J.-S. (2008) *Mostly harmless econometrics: An empiricist's companion*.

Princeton university press.

Angulo-Martínez, M. and Beguería, S. (2009) 'Estimating rainfall erosivity from daily precipitation records: A comparison among methods using data from the Ebro Basin (NE Spain)', *Journal of Hydrology*, 379(1–2), pp. 111–121.

Anttila-Hughes, J. and Hsiang, S. (2013) 'Destruction, disinvestment, and death: Economic and human losses following environmental disaster', *Available at SSRN 2220501*.

Apel, H. *et al.* (2015) 'Combined fluvial and pluvial urban flood hazard analysis: method development and application to Can Tho City, Mekong Delta, Vietnam', *Nat. Hazards Earth Syst. Sci. Discuss*, 3, pp. 4967–5013.

Arouri, M., Nguyen, C. and Youssef, A. Ben (2015) 'Natural disasters, household welfare, and resilience: evidence from rural Vietnam', *World development*, 70, pp. 59–77.

Arouri, M., Youssef, A. Ben and Nguyen, C. (2017) 'Does urbanization reduce rural poverty? Evidence from Vietnam', *Economic Modelling*, 60, pp. 253–270.

Asante, K. *et al.* (2019) 'The Ghana Randomized Air Pollution and Health Study (GRAPHS): a cluster randomized trial of Liquified petroleum gas (LPG) and efficient biomass cookstoves delivered during pregnancy', *Environmental Epidemiology*, 3, p. 296.

Ayongwa, G. C. *et al.* (2011) 'Organic matter and seed survival of Striga hermonthica–mechanisms for seed depletion in the soil', *Crop Protection*, 30(12), pp. 1594–1600.

Bacon, R., Bhattacharya, S. and Kojima, M. (2010) 'Expenditure of low-income households on energy: Evidence from Africa and Asia'.

Baez, J. E. *et al.* (2017) 'Gone with the Storm: Rainfall Shocks and Household Wellbeing in Guatemala', *The Journal of Development Studies*, 53(8), pp. 1253–1271. doi: 10.1080/00220388.2016.1224853.

Baez, J. E., Caruso, G. and Niu, C. (2020) 'Extreme weather and poverty risk: evidence from multiple shocks in Mozambique', *Economics of Disasters and Climate Change*, 4, pp. 103–127.

Baggio, J. A. and Papyrakis, E. (2010) 'Ethnic diversity, property rights, and natural resources', *The Developing Economies*, 48(4), pp. 473–495.

Di Baldassarre, G. *et al.* (2010) 'Flood fatalities in Africa: from diagnosis to mitigation', *Geophysical Research Letters*, 37(22).

Bandyopadhyay, S., Shyamsundar, P. and Baccini, A. (2011) 'Forests, biomass use and poverty in Malawi', *Ecological Economics*, 70(12), pp. 2461–2471.

Bandyopadhyay, S. and Skoufias, E. (2015) 'Rainfall variability, occupational choice, and welfare in rural Bangladesh', *Review of Economics of the Household*, 13, pp. 589–634.

Banerjee, L. (2010) 'Effects of flood on agricultural productivity in Bangladesh', *Oxford Development Studies*, 38(3), pp. 339–356.

Bangalore, M., Smith, A. and Veldkamp, T. (2019) 'Exposure to floods, climate change, and poverty in Vietnam', *Economics of Disasters and Climate Change*, 3, pp. 79–99.

Barbier, E. B. (2000) 'The economic linkages between rural poverty and land degradation: some evidence from Africa', *Agriculture, Ecosystems & Environment*, 82(1–3), pp. 355–370.

Barbier, E. B. (2010) 'Poverty, development, and environment', *Environment and Development Economics*, 15(6), pp. 635–660.

Barbier, E. B. (2012) 'Natural capital, ecological scarcity and rural poverty', *World Bank Policy Research Working Paper*, (6232).

Barbier, E. B. (2015) 'Climate change impacts on rural poverty in low-elevation coastal zones', *Estuarine, Coastal and Shelf Science*, 165, pp. A1–A13.

Barbier, E. B. and Hochard, J. P. (2016) 'Does land degradation increase poverty in developing countries?', *PloS one*, 11(5), p. e0152973.

Barbier, E. B. and Hochard, J. P. (2018) 'Land degradation and poverty', *Nature Sustainability*, 1(11), pp. 623–631.

Barrett, A. *et al.* (2013) 'Asset values: why are some households doing better than others', *Chars Livelihood Program*.

Barrett, C. B. (2010) 'Measuring food insecurity', *Science*, 327(5967), pp. 825–828.

Barrett, C. B. and Bevis, L. E. M. (2015a) 'The micronutrient deficiencies challenge in African Food Systems', *The fight against hunger and malnutrition: the role of food, agriculture, and targeted policies*, pp. 61–88.

Barrett, C. B. and Bevis, L. E. M. (2015b) 'The self-reinforcing feedback between low soil fertility and chronic poverty', *Nature Geoscience*, 8(12), pp. 907–912.

Barrett, C. B. and Swallow, B. M. (2006) 'Fractal poverty traps', *World development*, 34(1), pp. 1–15.

Barrett, C. B., Travis, A. J. and Dasgupta, P. (2011) 'On biodiversity conservation and poverty traps', *Proceedings of the National Academy of Sciences*, 108(34), pp. 13907–13912.

Barrios, S., Bertinelli, L. and Strobl, E. (2010) 'Trends in rainfall and economic growth in Africa: A neglected cause of the African growth tragedy', *The Review of Economics and Statistics*, 92(2), pp. 350–366.

Barron, M. and Torero, M. (2017) 'Household electrification and indoor air pollution', *Journal of Environmental Economics and Management*, 86, pp. 81–92.

Beegle, K. and Christiaensen, L. (2019) *Accelerating poverty reduction in Africa*. World Bank Publications.

BenYishay, A. *et al.* (2020) 'Gender gaps in technology diffusion', *Journal of Development Economics*, 143, p. 102380. doi: https://doi.org/10.1016/j.jdeveco.2019.102380.

Berman, N., Couttenier, M. and Soubeyran, R. (2021) 'Fertile ground for conflict', *Journal of the European Economic Association*, 19(1), pp. 82–127.

Berntsen, D. and Rubin, D. C. (2006) 'Flashbulb memories and posttraumatic stress reactions across the life span: age-related effects of the German occupation of Denmark during World War

II.', *Psychology and aging*, 21(1), p. 127.

Bertrand, M., Duflo, E. and Mullainathan, S. (2004) 'How much should we trust differences-in-differences estimates?', *The Quarterly journal of economics*, 119(1), pp. 249–275.

Bertrand, M. and Mullainathan, S. (2001) 'Do people mean what they say? Implications for subjective survey data', *American Economic Review*, 91(2), pp. 67–72.

Bhagwati, J. and Srinivasan, T. N. (2002) 'Trade and poverty in the poor countries', *American Economic Review*, 92(2), pp. 180–183.

Blakeslee, D., Fishman, R. and Srinivasan, V. (2020) 'Way down in the hole: Adaptation to long-term water loss in rural India', *American Economic Review*, 110(1), pp. 200–224.

Blattman, C. and Miguel, E. (2010) 'Civil war', *Journal of Economic literature*, 48(1), pp. 3–57.

Botzen, W. J. W., Deschenes, O. and Sanders, M. (2019) 'The economic impacts of natural disasters: A review of models and empirical studies', *Review of Environmental Economics and Policy*, 13(2), pp. 167–188.

Braubach, M. and Fairburn, J. (2010) 'Social inequities in environmental risks associated with housing and residential location—a review of evidence', *European journal of public health*, 20(1), pp. 36–42.

Brauer, M. *et al.* (2016) 'Ambient air pollution exposure estimation for the global burden of disease 2013', *Environmental science & technology*, 50(1), pp. 79–88.

von Braun, J. *et al.* (2013) *The economics of land degradation*. ZEF Working paper series.

Brenneman, A. and Kerf, M. (2002) 'Infrastructure & Poverty Linkages: A Literature Review (The World Bank Collection)'.

Bryden, K. M., Bruce, N. and Peck, M. (2015) 'WHO indoor air quality guidelines: household fuel combustion', *Geneva: World Health Organization*.

Budya, H. and Arofat, M. Y. (2011) 'Providing cleaner energy access in Indonesia through the megaproject of kerosene conversion to LPG', *Energy Policy*, 39(12), pp. 7575–7586.

Bui, A. T. *et al.* (2014) 'The impact of natural disasters on household income, expenditure, poverty and inequality: evidence from Vietnam', *Applied Economics*, 46(15), pp. 1751–1766.

Burwen, J. and Levine, D. I. (2012) 'A rapid assessment randomized-controlled trial of improved cookstoves in rural Ghana', *Energy for Sustainable Development*, 16(3), pp. 328–338.

Carter, M. R. *et al.* (2007) 'Poverty traps and natural disasters in Ethiopia and Honduras', *World development*, 35(5), pp. 835–856.

Cavallo, E. A., Becerra, O. and Acevedo, L. (2022) 'The impact of natural disasters on economic growth', in *Handbook on the Economics of Disasters*. Edward Elgar Publishing, pp. 150–192.

Ceola, S., Laio, F. and Montanari, A. (2014) 'Satellite nighttime lights reveal increasing human exposure to floods worldwide', *Geophysical Research Letters*, 41(20), pp. 7184–7190.

Cheema, I. *et al.* (2015) 'Citizen's damage compensation Programme: impact evaluation report'. Oxford Policy Management, Oxford, UK.

Chen, A., Giese, M. and Chen, D. (2020) ‘Flood impact on Mainland Southeast Asia between 1985 and 2018—The role of tropical cyclones’, *Journal of Flood Risk Management*, 13(2), p. e12598.

Chillrud, S. N. *et al.* (2021) ‘The effect of clean cooking interventions on mother and child personal exposure to air pollution: results from the Ghana Randomized Air Pollution and Health Study (GRAPHS)’, *Journal of exposure science & environmental epidemiology*, 31(4), pp. 683–698.

Chinh, D. T. *et al.* (2017) ‘Flood loss models and risk analysis for private households in Can Tho City, Vietnam’, *Water*, 9(5), p. 313.

Van Chinh, D., JianCheng, L. and Trinh, B. T. K. (2014) ‘Estimating the Design Values of Sea Level Heights at Some Tidal Stations along the Coast of Vietnam’.

Chomitz, K. M. *et al.* (2007) ‘At Loggerheads? Agricultural expansion, poverty reduction and environment in the tropical forests. The World Bank’.

Christiaensen, L., De Weerdt, J. and Todo, Y. (2013) ‘Urbanization and poverty reduction: the role of rural diversification and secondary towns 1’, *Agricultural Economics*, 44(4–5), pp. 435–447.

CIESIN (2016) ‘Gridded Population of the World, Version 4 (GPWv4): Administrative Unit Center Points with Population Estimates’. Palisades, NY: NASA Socioeconomic Data and Applications Center (SEDAC).

Le Cotty, T., Maitre D’Hotel, E. and Ndiaye, M. (2017) ‘Transport costs and food price volatility in Africa’, *Journal of African Economies*, 26(5), pp. 625–654.

CRED (2023) *EM-DAT The International Disasters Database*. Available at: https://www.emdat.be/.

Crespo Cuaresma, J. *et al.* (2017) ‘Economic development and forest cover: evidence from satellite data’, *Scientific reports*, 7(1), pp. 1–8.

Cuaresma, J. C. and Heger, M. (2019) ‘Deforestation and economic development: Evidence from national borders’, *Land Use Policy*, 84, pp. e347–e353.

Cutter, S. L. (1996) ‘Vulnerability to environmental hazards’, *Progress in human geography*, 20(4), pp. 529–539.

Le Dang, H. *et al.* (2014) ‘Farmers’ perceptions of climate variability and barriers to adaptation: lessons learned from an exploratory study in Vietnam’, *Mitigation and adaptation strategies for global change*, 19, pp. 531–548.

Dartmouth Flood Observatory (2018) *Dartmouth Flood Observatory Database*. Available at: https://floodobservatory.colorado.edu/.

Dartmouth Flood Observatory (2021) *DFO Flood Observatory: Space-based Measurement, Mapping, and Modeling of Surface Water*. Available at: https://floodobservatory.colorado.edu/index.html.

Dasgupta, S. *et al.* (2005) ‘Where is the poverty–environment nexus? Evidence from Cambodia, Lao PDR, and Vietnam’, *World Development*, 33(4), pp. 617–638.

Dercon, S., Hoddinott, J. and Woldehanna, T. (2005) ‘Shocks and consumption in 15 Ethiopian villages, 1999–2004’, *Journal of African economies*, 14(4), pp. 559–585.

DHI, M. (2011) ‘Mike 11-A modelling system for rivers and channels’, *Short Introduction and Tutorial, Horsholm, Denmark*.

Dillon, A. and Mensah, E. R. (2021) ‘Respondent Biases in Household Surveys’, *Global Poverty Research Lab Working Paper*, (21–103).

Duraiappah, A. K. (1998) ‘Poverty and environmental degradation: a review and analysis of the nexus’, *World development*, 26(12), pp. 2169–2179.

Ebenstein, A. *et al.* (2015) ‘Growth, pollution, and life expectancy: China from 1991–2012’, *American Economic Review*, 105(5), pp. 226–231.

Echendu, A. J. (2020) ‘The impact of flooding on Nigeria’s sustainable development goals (SDGs)’, *Ecosystem Health and Sustainability*, 6(1), p. 1791735. doi: 10.1080/20964129.2020.1791735.

Epstein, M. B. *et al.* (2013) ‘Household fuels, low birth weight, and neonatal death in India: the separate impacts of biomass, kerosene, and coal’, *International journal of hygiene and environmental health*, 216(5), pp. 523–532.

Erman, A. *et al.* (2018) *The road to recovery: the role of poverty in the exposure, vulnerability and resilience to floods in Accra*. The World Bank.

Erman, A. *et al.* (2019) ‘Flooded in Dar Es Salaam–the role of poverty in disaster risk’, *Policy Research Working Paper*, 8976.

Erman, A. *et al.* (2020) ‘The road to recovery the role of poverty in the exposure, vulnerability and resilience to floods in Accra’, *Economics of Disasters and Climate Change*, 4, pp. 171–193.

Eswaran, H. *et al.* (1997) ‘An assessment of the soil resources of Africa in relation to productivity’, *Geoderma*, 77(1), pp. 1–18.

Fanta, V., Šálek, M. and Sklenicka, P. (2019) ‘How long do floods throughout the millennium remain in the collective memory?’, *Nature communications*, 10(1), pp. 1–9.

Fathom (2023) *Flood Hazard Data & Maps*. Available at: https://www.fathom.global/product/flood-hazard-data-maps/.

Fay, M. (2005) *The urban poor in Latin America*. World Bank Publications.

Federal Government of Nigeria (2012) *NIGERIA Post-Disaster Needs Assessment 2012 Floods*. Available at: https://www.gfdrr.org/sites/default/files/publication/pda-2012-nigeria.pdf.

Federal Republic of Nigeria and National Bureau of Statistics (2016) *Basic Information Document Nigeria General Household Survey–Panel 2010 - 2011*.

Felbermayr, G. *et al.* (2022) ‘The economic impact of weather anomalies’, *World Development*, 151, p. 105745.

FIM (2013) ‘Final Report, Volume 2, Integrated Flood Management Strategy’. Ho Chi Minh City Flood and Inundation Management.

Forouzanfar, M. H. *et al.* (2016) ‘Global, regional, and national comparative risk assessment of 79 behavioural, environmental and occupational, and metabolic risks or clusters of risks, 1990–2015: a systematic analysis for the Global Burden of Disease Study 2015’, *The lancet*, 388(10053), pp.

1659–1724.

Fox, S. (2014) ‘The political economy of slums: Theory and evidence from Sub-Saharan Africa’, *World Development*, 54, pp. 191–203.

Fraval, S. *et al.* (2019) ‘Making the most of imperfect data: a critical evaluation of standard information collected in farm household surveys’, *Experimental Agriculture*, 55(2), pp. 230–250.

Fu, B. J. *et al.* (2005) ‘Assessment of soil erosion at large watershed scale using RUSLE and GIS: a case study in the Loess Plateau of China’, *Land degradation & development*, 16(1), pp. 73–85.

Galiani, S. and Schargrodsky, E. (2010) ‘Property rights for the poor: Effects of land titling’, *Journal of Public Economics*, 94(9–10), pp. 700–729.

Garrity, D. P. *et al.* (2010) ‘Evergreen Agriculture: a robust approach to sustainable food security in Africa’, *Food security*, 2, pp. 197–214.

Gennaioli, N. *et al.* (2013) ‘Human capital and regional development’, *The Quarterly journal of economics*, 128(1), pp. 105–164.

Gentle, P. *et al.* (2014) ‘Differential impacts of climate change on communities in the middle hills region of Nepal’, *Natural hazards*, 74, pp. 815–836.

Geographic Information Science and Technology (2015) ‘Landscan Population Data’. Oak Ridge National Laboratory.

Gerber, N., Nkonya, E. and von Braun, J. (2014) ‘Land degradation, poverty and marginality’, *Marginality: Addressing the nexus of poverty, exclusion and ecology*, pp. 181–202.

GFDRR (2015) ‘Country Profile - Vietnam’. Global Facility for Disaster Reduction and Recovery, World Bank Group. Available at: https://www.gfdrr.org/sites/gfdrr/files/region/VN.pdf.

Gibson, J. and Rozelle, S. (2003) ‘Poverty and access to roads in Papua New Guinea’, *Economic development and cultural change*, 52(1), pp. 159–185.

Global Forest Watch (2023) *Global Forest Watch Forest Monitoring*. Available at: https://www.globalforestwatch.org/.

Global Subsidies Initiative (2014) *Indonesia Energy Subsidy Review: A biannual survey of energy subsidy policies*.

Global Subsidies Initiative (2015) *Indonesia Energy Subsidy Review: A biannual survey of energy subsidy policies*.

Goeschl, T. and Managi, S. (2019) ‘Public in-kind relief and private self-insurance’, *Economics of Disasters and Climate Change*, 3, pp. 3–21.

González, F. A. I. (2022) ‘Natural Disasters and Economic Growth: a synthesis of empirical evidence’, *Nóesis. Revista de ciencias sociales*, 31(61), pp. 155–173.

Goodhand, J. (2001) ‘Violent conflict, poverty and chronic poverty’, *Chronic Poverty Research Centre Working Paper*, (6).

Gordon, S. B. *et al.* (2014) ‘Respiratory risks from household air pollution in low and middle income countries’, *The Lancet Respiratory Medicine*, 2(10), pp. 823–860.

Gould, C. F. *et al.* (2018) 'Government policy, clean fuel access, and persistent fuel stacking in Ecuador', *Energy for sustainable development*, 46, pp. 111–122.

Grainger, C. A. and Costello, C. J. (2014) 'Capitalizing property rights insecurity in natural resource assets', *Journal of Environmental Economics and Management*, 67(2), pp. 224–240.

Gray, L. C. and Moseley, W. G. (2005) 'A geographical perspective on poverty–environment interactions', *Geographical Journal*, 171(1), pp. 9–23.

Groom, B. *et al.* (2010) 'Relaxing rural constraints: a 'win-win'policy for poverty and environment in China?', *Oxford Economic Papers*, 62(1), pp. 132–156.

Guiteras, R., Jina, A. and Mobarak, A. M. (2015) 'Satellites, self-reports, and submersion: exposure to floods in Bangladesh', *American Economic Review*, 105(5), pp. 232–236.

Habitat for Humanity (2008) 'Mapping Urban Poverty in Ho Chi Minh City'. Available at: http://content.yudu.com/Library/A1gwji/MappingUrbanPovertyi/resources/8.htm.

Hagen-Zanker, J., Mallett, R. and Slater, R. (2015) 'The First Round of the SLRC Panel Survey: The Process Paper.', *Secure Livelihoods Research Consortium, Overseas Development Institute, London. https://assets. publishin g. service. gov. uk/media/57a0898f40f0b652dd0 002ae/SLRC-Panel-Survey-Process-Paper. pdf*.

Hainmueller, J. and Hangartner, D. (2019) 'Does direct democracy hurt immigrant minorities? Evidence from naturalization decisions in Switzerland', *American Journal of Political Science*, 63(3), pp. 530–547.

Hakam, D. F. *et al.* (2022) 'Mega conversion from LPG to induction stove to achieve Indonesia's clean energy transition', *Energy Strategy Reviews*, 41, p. 100856.

Hallegatte, S. *et al.* (2012) 'Investment decision making under deep uncertainty--application to climate change', *World Bank Policy Research Working Paper*, (6193).

Hallegatte, S., Bangalore, M., *et al.* (2016) *Shock Waves: Managing the Impacts of Climate Change on Poverty*. World Bank Publications.

Hallegatte, S., Vogt-Schilb, A., *et al.* (2016) *Unbreakable: building the resilience of the poor in the face of natural disasters*. World Bank Publications.

Hallegatte, S. *et al.* (2020) 'From poverty to disaster and back: A review of the literature', *Economics of Disasters and Climate Change*, 4, pp. 223–247.

Hanna, R., Duflo, E. and Greenstone, M. (2016) 'Up in smoke: the influence of household behavior on the long-run impact of improved cooking stoves', *American Economic Journal: Economic Policy*, 8(1), pp. 80–114.

Hansen, M. C. *et al.* (2013) 'High-resolution global maps of 21st-century forest cover change', *science*, 342(6160), pp. 850–853.

Harber, C. (2002) 'Education, democracy and poverty reduction in Africa', *Comparative education*, 38(3), pp. 267–276.

Harris, I. *et al.* (2014) 'Updated high-resolution grids of monthly climatic observations–the CRU TS3. 10 Dataset', *International journal of climatology*, 34(3), pp. 623–642.

Heltberg, R., Oviedo, A. M. and Talukdar, F. (2015) 'What do household surveys really tell us about risk, shocks, and risk management in the developing world?', *The journal of development studies*, 51(3), pp. 209–225.

Hernando, D. and Romana, M. G. (2015) 'Estimating the rainfall erosivity factor from monthly precipitation data in the Madrid Region (Spain)', *Journal of Hydrology and Hydromechanics*, 63(1), p. 55.

Hersher, R. (2022) *Climate change likely helped cause deadly Pakistan floods, scientists find*, *NPR*. Available at: https://www.npr.org/2022/09/19/1123798981/climate-change-likely-helped-cause-deadly-pakistan-floods-scientists-find.

Hiederer, R. and Köchy, M. (2011) 'Global soil organic carbon estimates and the harmonized world soil database', *EUR*, 79(25225), pp. 10–2788.

Hirabayashi, Y. *et al.* (2013) 'Global flood risk under climate change', *Nature climate change*, 3(9), pp. 816–821.

Hirst, W. *et al.* (2015) 'A ten-year follow-up of a study of memory for the attack of September 11, 2001: Flashbulb memories and memories for flashbulb events.', *Journal of Experimental Psychology: General*, 144(3), p. 604.

Horsley, W. (2004) 'Vietnam's Slum Dwellers'. BBC News.

Hulme, M. (1992) 'A 1951–80 global land precipitation climatology for the evaluation of general circulation models', *Climate dynamics*, 7, pp. 57–72.

Hulme, M. (1994) 'Validation of large-scale precipitation fields in general circulation models', in *Global precipitations and climate change*. Springer, pp. 387–405.

Hulme, M., Osborn, T. J. and Johns, T. C. (1998) 'Precipitation sensitivity to global warming: Comparison of observations with HadCM2 simulations', *Geophysical research letters*, 25(17), pp. 3379–3382.

Husby, T. and Hofkes, M. (2015) 'Loss Aversion on the Housing Market and Capitalisation of Flood Risk', *EAERE (European Association of Environmental and Resource Economists), Helsinki, Finland, June*, pp. 24–27.

Imelda (2018) 'The response of consumption to fuel switching: Panel data estimates'.

Imelda (2020) 'Cooking that kills: Cleaner energy access, indoor air pollution, and health', *Journal of Development Economics*, 147, p. 102548. doi: https://doi.org/10.1016/j.jdeveco.2020.102548.

IMHEN and UNDP (2015) 'Summary for Policy Makers'. Natural Resources and Environment Publishing House, Hanoi.

IPCC (2007) *Summary for Policymakers*. Cambridge University Press, Cambridge, UK and New York, USA. Available at: http://www.ipcc.ch/pdf/assessment-report/ar4/syr/ar4_syr_spm.pdf.

IPCC (2014) *Climate Change 2014: Impacts, Adaptation, and Vulnerability. Part B: Regional Aspects*. Cambridge University Press, Cambridge, UK and New York, USA.

IPCC (2018) *Global warming of 1.5°C. An IPCC Special Report on the impacts of global warming of 1.5°C above pre-industrial levels and related global greenhouse gas emission pathways, in the* *context of strengthening the global response to the threat of climate change,*.

IPCC (2023) 'AR6 Synthesis Report: Climate Change 2023'. Available at: https://www.ipcc.ch/report/sixth-assessment-report-cycle/.

ISM; SINFONICA (2015) 'Users' manual for handling resampled micro data of Vietnamese Household Living Standard Survey (VHLSS)'', *Institute of Statistical Mathematics and Statistical Information Institute for Consulting and Analysis*.

Jack, D. W. *et al.* (2021) 'A cluster randomised trial of cookstove interventions to improve infant health in Ghana', *BMJ Global Health*, 6(8), p. e005599.

Jacoby, H. G. (2000) 'Access to markets and the benefits of rural roads', *The economic journal*, 110(465), pp. 713–737.

Jacoby, H. G. and Minten, B. (2009) 'On measuring the benefits of lower transport costs', *Journal of Development Economics*, 89(1), pp. 28–38.

De Janvry, A. *et al.* (2006) 'Can conditional cash transfer programs serve as safety nets in keeping children at school and from working when exposed to shocks?', *Journal of development economics*, 79(2), pp. 349–373.

Al Jazeera (2022) *Nigeria flood death toll tops 600 as thousands evacuated*. Available at: https://www.aljazeera.com/news/2022/10/17/nigeria-flood-death-toll-rises-as-thousands-evacuated.

Jongman, B. *et al.* (2014) 'Increasing stress on disaster-risk finance due to large floods', *Nature Climate Change*, 4(4), pp. 264–268.

Kakietek, J. J. *et al.* (2022) 'Foregone healthcare during the COVID-19 pandemic: early survey estimates from 39 low-and middle-income countries', *Health Policy and Planning*, 37(6), pp. 771–778.

Kang, G. Z. *et al.* (2013) 'Exogenous salicylic acid enhances wheat drought tolerance by influence on the expression of genes related to ascorbate-glutathione cycle', *Biologia plantarum*, 57(4), pp. 718–724.

Karim, A. and Noy, I. (2016) 'Poverty and natural disasters: a regression meta-analysis', *Review of Economics and Institutions*, 7(2), p. 26.

Katz, J. *et al.* (2020) 'Impact of improved biomass and liquid petroleum gas stoves on birth outcomes in rural Nepal: results of 2 randomized trials', *Global Health: Science and Practice*, 8(3), pp. 372–382.

Kazianga, H. and Udry, C. (2006) 'Consumption smoothing? Livestock, insurance and drought in rural Burkina Faso', *Journal of Development economics*, 79(2), pp. 413–446.

Khan, Shaheen Rafi and Khan, Shahrukh Rafi (2009) 'Assessing poverty–deforestation links: Evidence from Swat, Pakistan', *Ecological Economics*, 68(10), pp. 2607–2618.

Khandker, S. R., Bakht, Z. and Koolwal, G. B. (2009) 'The poverty impact of rural roads: Evidence from Bangladesh', *Economic development and cultural change*, 57(4), pp. 685–722.

Klomp, J. (2016) 'Economic development and natural disasters: A satellite data analysis', *Global*

Environmental Change, 36, pp. 67–88.

Klomp, J. and Valckx, K. (2014) 'Natural disasters and economic growth: A meta-analysis', *Global Environmental Change*, 26, pp. 183–195.

Kocornik-Mina, A. *et al.* (2020) 'Flooded cities', *American Economic Journal: Applied Economics*, 12(2), pp. 35–66.

Koks, E. E. *et al.* (2015) 'Combining hazard, exposure and social vulnerability to provide lessons for flood risk management', *Environmental science & policy*, 47, pp. 42–52.

Koren, O. (2018) 'Food abundance and violent conflict in Africa', *American Journal of Agricultural Economics*, 100(4), pp. 981–1006.

Kozel, V. (2014) *Well begun but not yet done: Progress and emerging challenges for poverty reduction in Vietnam*. World Bank Publications.

Krishna, A. *et al.* (2006) 'Escaping poverty and becoming poor in 36 villages of Central and Western Uganda', *The Journal of Development Studies*, 42(2), pp. 346–370.

Krishnapriya, P. P. *et al.* (2021) 'Do improved cookstoves save time and improve gender outcomes? Evidence from six developing countries', *Energy Economics*, 102, p. 105456.

Kundzewicz, Z. W., Pińskwar, I. and Brakenridge, G. R. (2013) 'Large floods in Europe, 1985–2009', *Hydrological Sciences Journal*, 58(1), pp. 1–7.

Kunreuther, H. C., Pauly, M. V and McMorrow, S. (2013) *Insurance and behavioral economics: Improving decisions in the most misunderstood industry*. Cambridge University Press.

Kurt, O. K., Zhang, J. and Pinkerton, K. E. (2016) 'Pulmonary health effects of air pollution', *Current opinion in pulmonary medicine*, 22(2), p. 138.

Lal, R. (2004) 'Soil carbon sequestration impacts on global climate change and food security', *science*, 304(5677), pp. 1623–1627.

Lall, S. V and Deichmann, U. (2012) 'Density and disasters: economics of urban hazard risk', *The World Bank Research Observer*, 27(1), pp. 74–105.

Lambin, E. F. *et al.* (2001) 'The causes of land-use and land-cover change: moving beyond the myths', *Global environmental change*, 11(4), pp. 261–269.

Lanjouw, P. F., Marra, M. and Nguyen, C. V. (2013) 'Vietnam's evolving poverty map: patterns and implications for policy', *World Bank Policy Research Working Paper*, (6355).

Lasage, R. *et al.* (2014) 'Assessment of the effectiveness of flood adaptation strategies for HCMC', *Natural Hazards and Earth System Sciences*, 14(6), pp. 1441–1457.

Latifah, E. W. (2010) 'Analisis Persepsi, Sikap, dan Strategi koping keluarga miskin terkait program konversi minyak tanah ke LPG di kota Bogor'.

Laurance, W. F., Sayer, J. and Cassman, K. G. (2014) 'Agricultural expansion and its impacts on tropical nature', *Trends in ecology & evolution*, 29(2), pp. 107–116.

De Lay, S. (2011) 'Slums in Ho Chi Minh City, Vietnam'. Global Cities 2011. Available at: http://hochiminhcity2011.jimdo.com/slums/.

Lazzaroni, S. and van Bergeijk, P. A. G. (2014) 'Natural disasters' impact, factors of resilience and development: A meta-analysis of the macroeconomic literature', *Ecological Economics*, 107, pp. 333–346.

Li, Y. *et al.* (2013) 'Rural livelihoods and climate variability in Ningxia, Northwest China', *Climatic Change*, 119, pp. 891–904.

Liao, K.-H., Le, T. A. and Van Nguyen, K. (2016) 'Urban design principles for flood resilience: Learning from the ecological wisdom of living with floods in the Vietnamese Mekong Delta', *Landscape and Urban Planning*, 155, pp. 69–78.

Lima, R. C. de A. and Barbosa, A. V. B. (2019) 'Natural disasters, economic growth and spatial spillovers: Evidence from a flash flood in Brazil', *Papers in Regional Science*, 98(2), pp. 905–924.

Lin, N. *et al.* (2012) 'Physically based assessment of hurricane surge threat under climate change', *Nature Climate Change*, 2(6), pp. 462–467.

List, J. A. (2020) *Non est disputandum de generalizability? a glimpse into the external validity trial*. National Bureau of Economic Research.

Liu, Y. *et al.* (2014) 'Impacts of droughts on carbon sequestration by China's terrestrial ecosystems from 2000 to 2011', *Biogeosciences*, 11(10), pp. 2583–2599.

Louwagie, G. *et al.* (2011) 'The potential of European Union policies to address soil degradation in agriculture', *Land degradation & development*, 22(1), pp. 5–17.

Lowe, J. A. *et al.* (2010) 'Past and future changes in extreme sea levels and waves'.

Lufafa, A. *et al.* (2003) 'Prediction of soil erosion in a Lake Victoria basin catchment using a GIS-based Universal Soil Loss model', *Agricultural systems*, 76(3), pp. 883–894.

Luft, H. S. (1975) 'The impact of poor health on earnings', *The Review of Economics and Statistics*, pp. 43–57.

Luong, L. M. T. *et al.* (2017) 'The association between particulate air pollution and respiratory admissions among young children in Hanoi, Vietnam', *Science of the total environment*, 578, pp. 249–255.

Masters, W. A. *et al.* (2013) 'Urbanization and farm size in Asia and Africa: Implications for food security and agricultural research', *Global Food Security*, 2(3), pp. 156–165.

McCarthy, N. *et al.* (2018) 'Shelter from the storm? household-level impacts of, and responses to, the 2015 floods in Malawi', *Economics of disasters and climate change*, 2(3), pp. 237–258.

McCarthy, N. *et al.* (2021) 'Droughts and floods in Malawi: impacts on crop production and the performance of sustainable land management practices under weather extremes', *Environment and Development Economics*, 26(5–6), pp. 432–449.

McCracken, J. P. *et al.* (2007) 'Chimney stove intervention to reduce long-term wood smoke exposure lowers blood pressure among Guatemalan women', *Environmental health perspectives*, 115(7), pp. 996–1001.

McElwee, P. *et al.* (2017) 'Flood vulnerability among rural households in the Red River Delta of Vietnam: implications for future climate change risk and adaptation', *Natural Hazards*, 86, pp.

465–492.

Meles, T. H. (2020) ‘Impact of power outages on households in developing countries: Evidence from Ethiopia’, *Energy Economics*, 91, p. 104882.

Miguel, E. (2005) ‘Poverty and witch killing’, *The Review of Economic Studies*, 72(4), pp. 1153–1172.

Miguel, E., Satyanath, S. and Sergenti, E. (2004) ‘Economic shocks and civil conflict: An instrumental variables approach’, *Journal of political Economy*, 112(4), pp. 725–753.

Miller, G. and Mobarak, A. M. (2013) *Gender differences in preferences, intra-household externalities, and low demand for improved cookstoves*. National Bureau of Economic Research.

Ministry of Natural Resources and Environment (2009) *Climate Change, Sea Level Rise Scenarios for Vietnam*.

Mortimer, K. *et al.* (2017) ‘A cleaner burning biomass-fuelled cookstove intervention to prevent pneumonia in children under 5 years old in rural Malawi (the Cooking and Pneumonia Study): a cluster randomised controlled trial’, *The Lancet*, 389(10065), pp. 167–175.

Mueller, L. *et al.* (2011) ‘Assessing the productivity function of soils’, *Sustainable Agriculture Volume 2*, pp. 743–760.

Naipal, V. *et al.* (2015) ‘Improving the global applicability of the RUSLE model–adjustment of the topographical and rainfall erosivity factors’, *Geoscientific Model Development*, 8(9), pp. 2893–2913.

Narloch, U. and Bangalore, M. (2018) ‘The multifaceted relationship between environmental risks and poverty: new insights from Vietnam’, *Environment and Development Economics*, 23(3), pp. 298–327.

Narloch, U. G. (2016) ‘The varying income effects of weather variation: Initial insights from rural Vietnam’, *World Bank Policy Research Working Paper*, (7764).

Narloch, U. G. and Bangalore, M. (2016) ‘Environmental risks and poverty: analyzing geo-spatial and household data from Vietnam’, *World Bank Policy Research Working Paper*, (7763).

NASA (2022) *NRT Global Flood Mapping Product Description*. Available at: https://floodmap.modaps.eosdis.nasa.gov/proddesc.php.

NASA (2023) *MODIS Near Real-Time Global Flood Product*. Available at: https://www.earthdata.nasa.gov/learn/find-data/near-real-time/modis-nrt-global-flood-product.

Naz, S., Page, A. and Agho, K. E. (2017) ‘Household air pollution from use of cooking fuel and under-five mortality: The role of breastfeeding status and kitchen location in Pakistan’, *PLoS One*, 12(3), p. e0173256.

NCAR (2017) ‘The Climate Data Guide: CRU TS Gridded precipitation and other meteorological variables since 1901’.

Neal, J., Schumann, G. and Bates, P. (2012) ‘A subgrid channel model for simulating river hydraulics and floodplain inundation over large and data sparse areas’, *Water Resources Research*, 48(11).

Nigeria Bureau of Statistics (2023) *Nigeria Bureau of Statistics Food Price Dataset*. Available at: https://www.nigerianstat.gov.ng/.

NIMET (2012) *Seasonal rainfall prediction and socio-economic implications for Nigeria*.

Del Ninno, C. (2001) *The 1998 floods in Bangladesh: disaster impacts, household coping strategies, and response*. Intl Food Policy Res Inst.

Nkonya, E., Mirzabaev, A. and Von Braun, J. (2016) *Economics of land degradation and improvement–a global assessment for sustainable development*. Springer Nature.

Noy, I., Nguyen, C. and Patel, P. (2021) 'Floods and spillovers: Households after the 2011 great flood in Thailand', *Economic Development and Cultural Change*, 69(2), pp. 829–868.

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

Nzeribe, G. E., Nwokoye, E. and Ezenekwe, U. (2014) 'What Implications does 2012 Flood Disaster have on the Nigerian Economy?', *Association of Energy and Environmental Economics of Nigeria-Nigerian Journal of Energy and Environmental Economics*, 6(1), pp. 15–28.

OCHA (2012) *Nigeria: Floods Situation Report No. 2 (as of 15 November 2012)*.

Odjugo, P. A. O. (2009) 'Quantifying the cost of climate change impact in Nigeria: Emphasis on wind and rainstorms', *Journal of human ecology*, 28(2), pp. 93–101.

Okwi, P. O. *et al.* (2007) 'Spatial determinants of poverty in rural Kenya', *Proceedings of the National Academy of Sciences*, 104(43), pp. 16769–16774.

Oladokun, V. O. and Proverbs, D. (2016) 'Flood risk management in Nigeria: A review of the challenges and opportunities', *Flood Risk Management and Response*, 6(3), pp. 485–497.

Olaore, A. Y. and Aja, G. N. (2014) 'The Impact of Flooding on the Social Determinants of Health in Nigeria: A Case for North-South Institutional Collaboration to Address Climate Issues', *Developing Country Studies*, 4(22), pp. 6–13.

Olinto, P. *et al.* (2013) 'The state of the poor: Where are the poor, where is extreme poverty harder to end, and what is the current profile of the world's poor', *Economic premise*, 125(2), pp. 1–8.

Opondo, D. O. (2013) 'Erosive coping after the 2011 floods in Kenya', *International Journal of Global Warming*, 5(4), pp. 452–466.

Osumgborogwu, I. E. and Chibo, C. N. (2017) 'Environmental laws in Nigeria and occurrence of some geohazards: a review', *Asian Journal of Environment & Ecology*, pp. 1–12.

Panayotou, T. (1997) 'Demystifying the environmental Kuznets curve: turning a black box into a policy tool', *Environment and development economics*, 2(4), pp. 465–484.

Panayotou, T. (2016) 'Economic growth and the environment. The environment in anthropology'. New York University Press, New York.

Pasanen, T. *et al.* (2017) 'Poverty–environment nexus in the Lao PDR: Analysis of household survey data', *Development Policy Review*, 35(3), pp. 349–371.

Patankar, A. M. (2015) ‘The exposure, vulnerability, and ability to respond of poor households to recurrent floods in Mumbai’, *World Bank Policy Research Working Paper*, (7481).

Paxson, C. H. (1992) ‘Using weather variability to estimate the response of savings to transitory income in Thailand’, *The American Economic Review*, pp. 15–33.

Pellegrino, R. *et al.* (2005) ‘Interpretative strategies for lung function tests’, *European respiratory journal*, 26(5), pp. 948–968.

Perez-De-Rada, E. and Paz, D. (2008) ‘Análisis de la relación entre amenazas naturales y condiciones de vida: El Caso de Bolivia’, *ISDR/RBLAC Research Project on Disaster Risk and Poverty*.

Pertamina (2012) ‘Kerosene to LP gas conversion programme in Indonesia: A case study of domestic energy’, *Neuilly-Sur-Seine: WLPGA Publications*.

Petley, D. (2012) ‘Global patterns of loss of life from landslides’, *Geology*, 40(10), pp. 927–930.

Petley, D. N. (2010) ‘On the impact of climate change and population growth on the occurrence of fatal landslides in South, East and SE Asia’, in. Geological Society of London.

Pham, T. T. *et al.* (2012) ‘The context of REDD+ in Vietnam: drivers, agents and institutions.’, *CIFOR Occasional Paper*, (75).

Phillips, L. B., Hansen, A. J. and Flather, C. H. (2008) ‘Evaluating the species energy relationship with the newest measures of ecosystem energy: NDVI versus MODIS primary production’, *Remote Sensing of Environment*, 112(12), pp. 4381–4392.

Pitt, M. M., Rosenzweig, M. R. and Hassan, M. (2005) ‘Sharing the burden of disease: gender, the household division of labor and the health effects of indoor air pollution’, *CID Working Paper Series*.

Playán, E., Sagardoy, J. A. and Castillo, R. (2018) ‘Irrigation governance in developing countries: Current problems and solutions’, *Water*, 10(9), p. 1118.

Policelli, F. *et al.* (2017) ‘The NASA global flood mapping system’, *Remote sensing of hydrological extremes*, pp. 47–63.

PUMA (2013) ‘Interim Report Version 1.0’. Platform for Urban Management and Analysis (PUMA) Software Development and Satellite Imagery Processing Consultants.

Quinn, A. K. *et al.* (2018) ‘An analysis of efforts to scale up clean household energy for cooking around the world’, *Energy for Sustainable Development*, 46, pp. 1–10.

Rabbani, G., Rahman, A. and Mainuddin, K. (2013) ‘Salinity-induced loss and damage to farming households in coastal Bangladesh’, *International Journal of Global Warming*, 5(4), pp. 400–415.

Ranger, N. *et al.* (2011) ‘An assessment of the potential impact of climate change on flood risk in Mumbai’, *Climatic change*, 104, pp. 139–167.

Ravallion, M. (2001) ‘Growth, inequality and poverty: looking beyond averages’, *World development*, 29(11), pp. 1803–1815.

Reardon, T. and Vosti, S. A. (1995) ‘Links between rural poverty and the environment in developing countries: asset categories and investment poverty', *World development*, 23(9), pp. 1495–1506.

De Ree, J. *et al.* (2018) 'Double for nothing? Experimental evidence on an unconditional teacher salary increase in Indonesia', *The Quarterly Journal of Economics*, 133(2), pp. 993–1039.

Rentschler, J. and Kornejew, M. (2017) 'Energy price variation and competitiveness: Firm level evidence from Indonesia', *Energy Economics*, 67, pp. 242–254.

Rentschler, J., Salhab, M. and Jafino, B. A. (2022) 'Flood exposure and poverty in 188 countries', *Nature communications*, 13(1), p. 3527.

Rodriguez-Oreggia, E. *et al.* (2013) 'Natural disasters, human development and poverty at the municipal level in Mexico', *The Journal of Development Studies*, 49(3), pp. 442–455.

Rozenberg, J. and Hallegatte, S. (2016) 'Modeling the impacts of climate change on future Vietnamese households: a micro-simulation approach', *World Bank Policy Research Working Paper*, (7766).

Russell, A., Locatelli, B. and Pramova, E. (2012) 'Using forests to enhance resilience to climate change: What do we know about how forests can contribute to adaptation?'

Safir, A., Piza, S. F. and Skoufias, E. (2013) 'Disquiet on the weather front: The welfare impacts of climatic variability in the rural Philippines', *World Bank Policy Research Working Paper*, (6579).

Sager, L. (2019) 'Estimating the effect of air pollution on road safety using atmospheric temperature inversions', *Journal of Environmental Economics and Management*, 98, p. 102250.

Sampson, C. C. *et al.* (2015) 'A high-resolution global flood hazard model', *Water resources research*, 51(9), pp. 7358–7381.

Sanchez, P. A. *et al.* (1997) 'Soil fertility replenishment in Africa: an investment in natural resource capital', *Replenishing soil fertility in Africa*, 51, pp. 1–46.

Sarsons, H. (2015) 'Rainfall and conflict: A cautionary tale', *Journal of development Economics*, 115, pp. 62–72.

Scherr, S. J. (2000) 'A downward spiral? Research evidence on the relationship between poverty and natural resource degradation', *Food policy*, 25(4), pp. 479–498.

Schippers, P. *et al.* (2015) 'Tree growth variation in the tropical forest: understanding effects of temperature, rainfall and CO 2', *Global change biology*, 21(7), pp. 2749–2761.

Scussolini, P. *et al.* (2015) 'FLOPROS: an evolving global database of flood protection standards', *Nat. Hazards Earth Syst. Sci. Discuss*, 3, pp. 7275–7309.

Selod, H. and Shilpi, F. (2021) 'Rural-urban migration in developing countries: Lessons from the literature', *Regional Science and Urban Economics*, 91, p. 103713.

Seneviratne, S. *et al.* (2012) 'Changes in climate extremes and their impacts on the natural physical environment'.

Shiru, M. S. *et al.* (2019) 'Changing characteristics of meteorological droughts in Nigeria during 1901–2010', *Atmospheric Research*, 223, pp. 60–73.

Sidle, R. C. *et al.* (2006) ‘Erosion processes in steep terrain—Truths, myths, and uncertainties related to forest management in Southeast Asia’, *Forest ecology and management*, 224(1–2), pp. 199–225.

Sietz, D., Mamani Choque, S. E. and Lüdeke, M. K. B. (2012) ‘Typical patterns of smallholder vulnerability to weather extremes with regard to food security in the Peruvian Altiplano’, *Regional Environmental Change*, 12, pp. 489–505.

Silwal, A. R. and McKay, A. (2015) ‘The impact of cooking with firewood on respiratory health: Evidence from Indonesia’, *The Journal of Development Studies*, 51(12), pp. 1619–1633.

Sissoko, K. *et al.* (2011) ‘Agriculture, livelihoods and climate change in the West African Sahel’, *Regional Environmental Change*, 11, pp. 119–125.

Skoufias, E. *et al.* (2001) ‘Conditional cash transfers and their impact on child work and schooling: Evidence from the progresa program in mexico [with comments]’, *Economia*, 2(1), pp. 45–96.

Slayback, D. (2023) *LANCE MODIS Flood Product User Guide*.

Sloan, S. and Sayer, J. A. (2015) ‘Forest Resources Assessment of 2015 shows positive global trends but forest loss and degradation persist in poor tropical countries’, *Forest Ecology and Management*, 352, pp. 134–145.

Smith, A. *et al.* (2014) ‘Comparing ensemble projections of flooding against flood estimation by continuous simulation’, *Journal of Hydrology*, 511, pp. 205–219.

Smith, A., Sampson, C. and Bates, P. (2015) ‘Regional flood frequency analysis at the global scale’, *Water Resources Research*, 51(1), pp. 539–553.

Smith, K. R. *et al.* (2011) ‘Effect of reduction in household air pollution on childhood pneumonia in Guatemala (RESPIRE): a randomised controlled trial’, *The Lancet*, 378(9804), pp. 1717–1726.

Smith, K. R. *et al.* (2014) ‘Millions dead: how do we know and what does it mean? Methods used in the comparative risk assessment of household air pollution’, *Annual review of public health*, 35, pp. 185–206.

Sodhi, N. S. *et al.* (2010) ‘Conserving Southeast Asian forest biodiversity in human-modified landscapes’, *Biological Conservation*, 143(10), pp. 2375–2384.

Somanathan, E. *et al.* (2022) *Electric stoves as a solution for household air pollution: Evidence from rural India*. Indian Statistical Institute, Delhi, Economics and Planning Unit.

Stern, D. I. (2004) ‘The rise and fall of the environmental Kuznets curve’, *World development*, 32(8), pp. 1419–1439.

Strauss, J. and Witoelar, F. (2022) ‘Indonesia family life survey’, in *Encyclopedia of gerontology and population aging*. Springer, pp. 2600–2605.

Suich, H., Howe, C. and Mace, G. (2015) ‘Ecosystem services and poverty alleviation: a review of the empirical links’, *Ecosystem Services*, 12, pp. 137–147.

Tanoue, M., Hirabayashi, Y. and Ikeuchi, H. (2016) ‘Global-scale river flood vulnerability in the last 50 years’, *Scientific reports*, 6(1), p. 36021.

Tansley, G. *et al.* (2017) 'The correlation between poverty and access to essential surgical care in Ghana: a geospatial analysis', *World journal of surgery*, 41, pp. 639–643.

Tellman, B. *et al.* (2020) 'Satellite observations indicate increasing proportion of population exposed to floods'.

Tellman, B. (2023) *Understanding Flood Risk from Space: Opportunities to adapt to changing risk and catalyze climate justice*.

Tenzing, J. D. (2020) 'Integrating social protection and climate change adaptation: A review', *Wiley Interdisciplinary Reviews: Climate Change*, 11(2), p. e626.

Thoday, K. *et al.* (2018) 'The Mega Conversion Program from Kerosene to LPG in Indonesia: lessons learned and recommendations for future clean cooking energy expansion.', *Energy for sustainable development : the journal of the International Energy Initiative*, 46, pp. 71–81. doi: 10.1016/j.esd.2018.05.011.

Thomas, T. *et al.* (2010) 'Natural disasters and household welfare: evidence from Vietnam', *World Bank Policy Research Working Paper*, (5491).

Tiague, B. D. (2021) *Floods, Agricultural Productivity, and Household Welfare: Evidence from Tanzania*.

Toft, L., Beaton, C. and Lontoh, L. (2016) *International experiences with LPG subsidy reform*. International Institute for Sustainable Development.

Toman, M. and Bluffstone, R. (2017) 'Challenges in assessing the costs of household cooking energy in lower-income countries'.

Trinh, T. Q. *et al.* (2018) 'Determinants of farmers' adaptation to climate change in agricultural production in the central region of Vietnam', *Land use policy*, 70, pp. 224–231.

Tschakert, P. (2007) 'Views from the vulnerable: understanding climatic and other stressors in the Sahel', *Global Environmental Change*, 17(3–4), pp. 381–396.

Tucker, J. *et al.* (2015) 'Social vulnerability in three high-poverty climate change hot spots: What does the climate change literature tell us?', *Regional Environmental Change*, 15, pp. 783–800.

UC Louvain, CRED and USAID (2022) *2021 Disasters in numbers*. Available at: https://reliefweb.int/report/world/2021-disasters-numbers.

Umar, N. and Gray, A. (2022) 'Flooding in Nigeria: a review of its occurrence and impacts and approaches to modelling flood data', *International Journal of Environmental Studies*, pp. 1–22.

Unah, L. (2017) 'Flood-ridden Nigeria farmers need more help adapting to climate change', *The New Humanitarian*, 3 May. Available at: https://www.thenewhumanitarian.org/feature/2017/05/03/flood-ridden-nigeria-farmers-need-more-help-adapting-climate-change.

USDA (1999) 'Soil Taxonomy: A Basic System of Soil Classification for Making and Interpreting Soil Surveys.' Washington, DC: United States Department of Agriculture.

Del Valle, A. *et al.* (2018) 'The short-term economic impact of tropical cyclones: Satellite evidence from Guangdong province', *Economics of Disasters and Climate Change*, 2, pp. 225–235.

Verma, A. P. and Imelda (2023) ‘Clean Energy Access: Gender Disparity, Health and Labour Supply’, *The Economic Journal*, 133(650), pp. 845–871.

Vezina, K., Bonn, F. and Van, C. P. (2006) ‘Agricultural land-use patterns and soil erosion vulnerability of watershed units in Vietnam’s northern highlands’, *Landscape Ecology*, 21, pp. 1311–1325.

Vlam, M. *et al.* (2014) ‘Temperature and rainfall strongly drive temporal growth variation in Asian tropical forest trees’, *Oecologia*, 174, pp. 1449–1461.

Vu, Q. M. *et al.* (2014) ‘Socio-economic and biophysical determinants of land degradation in Vietnam: An integrated causal analysis at the national level’, *Land use policy*, 36, pp. 605–617.

Vu, Q. M., Le, Q. B. and Vlek, P. L. G. (2014) ‘Hotspots of human-induced biomass productivity decline and their social–ecological types toward supporting national policy and local studies on combating land degradation’, *Global and Planetary Change*, 121, pp. 64–77.

Wantchekon, L. and Stanig, P. (2015) ‘The curse of good soil? Land fertility, roads, and rural poverty in Africa’, *Unpublished manuscript*.

Watmough, G. R. *et al.* (2016) ‘Understanding the evidence base for poverty–environment relationships using remotely sensed satellite data: an example from Assam, India’, *World Development*, 78, pp. 188–203.

Wing, O. E. J. *et al.* (2017) ‘Validation of a 30 m resolution flood hazard model of the conterminous U nited S tates’, *Water Resources Research*, 53(9), pp. 7968–7986.

Winsemius, H. C. *et al.* (2018) ‘Disaster risk, climate change, and poverty: assessing the global exposure of poor people to floods and droughts’. The World Bank.

Wiratmaja, I. (2016) *Update on fossil fuel subsidies in Indonesia*.

Wizor, C. H. and Week, D. A. (2014) ‘Impact of the 2012 Nigeria flood on emergent cities of Nigeria: The case of Yenagoa, Bayelsa State’, *Civil and Environmental Research*, 6(5), pp. 31–40.

Wodon, Q. *et al.* (2014) *Climate change and migration: evidence from the Middle East and North Africa*. World Bank Publications.

Wong, M. T. F. and Asseng, S. (2006) ‘Determining the causes of spatial and temporal variability of wheat yields at sub-field scale using a new method of upscaling a crop model’, *Plant and Soil*, 283, pp. 203–215.

World Bank (2009) ‘Reshaping Economic Geography’.

World Bank (2010) ‘Economics of Adaptation to Climate Change in Vietnam’.

World Bank (2014) ‘Turn Down the Heat: Climate Extremes, Regional Impacts, and the Case for Resilience’.

World Bank (2015a) *Another Nargis Strikes Everyday: Post-Nargis Social Impacts Monitoring Five Years On*. Washington, DC.

World Bank (2015b) ‘Platform for Urban Management and Analysis’. Available at: http://puma.worldbank.org/.

World Bank (2016a) ‘Climate Change and Poverty in Vietnam’. World Bank Internal Document.

World Bank (2016b) *Poverty and shared prosperity 2016: taking on inequality*. The World Bank.

World Bank (2021) *World Development Indicators*.

World Bank (2023a) ‘Hidden Dimensions Of Poverty Dataset (HDD)’. Available at: https://datacatalog.worldbank.org/search/dataset/0021031/Hidden-Dimensions-of-Poverty-Dataset--HDD-.

World Bank (2023b) *Indonesia Database For Policy And Economic Research*. Available at: https://datacatalog.worldbank.org/search/dataset/0041056.

World Bank and Australian AID (2014) *‘WHERE ARE WE DURING FLOODING?’ A Qualitative Assessment of Poverty and Social Impacts of Flooding in Selected Neighborhoods of HCMC*.

World Bank and GFDRR (2011) ‘Vulnerability, Risk Reduction, and Adaptation to Climate Change - Vietnam’.

World Health Organization (2022) *Household air pollution*. Available at: https://www.who.int/news-room/fact-sheets/detail/household-air-pollution-and-health.

World Population Review (2023) *Ho Chi Minh City Population 2023*. Available at: https://worldpopulationreview.com/world-cities/ho-chi-minh-city-population.

Wunder, S. (2015) ‘Revisiting the concept of payments for environmental services’, *Ecological economics*, 117, pp. 234–243.

Wunder, S., Noack, F. and Angelsen, A. (2018) ‘Climate, crops, and forests: a pan-tropical analysis of household income generation’, *Environment and Development Economics*, 23(3), pp. 279–297.

Yan, H. *et al.* (2014) ‘Development of a remotely sensing seasonal vegetation-based Palmer drought severity index and its application of global drought monitoring over 1982–2011’, *Journal of Geophysical Research: Atmospheres*, 119(15), pp. 9419–9440.

Yang, D. and Choi, H. (2007) ‘Are remittances insurance? Evidence from rainfall shocks in the Philippines’, *The World Bank Economic Review*, 21(2), pp. 219–248.

You, L. *et al.* (2011) ‘What is the irrigation potential for Africa? A combined biophysical and socioeconomic approach’, *Food Policy*, 36(6), pp. 770–782.

Zimmerman, F. J. and Carter, M. R. (2003) ‘Asset smoothing, consumption smoothing and the reproduction of inequality under risk and subsistence constraints’, *Journal of Development Economics*, 71(2), pp. 233–260.

Zuazo, V. H. D. and Pleguezuelo, C. R. R. (2009) ‘Soil-erosion and runoff prevention by plant covers: a review’, *Sustainable agriculture*, pp. 785–811.

Zulu, L. C. and Richardson, R. B. (2013) ‘Charcoal, livelihoods, and poverty reduction: Evidence from sub-Saharan Africa’, *Energy for Sustainable Development*, 17(2), pp. 127–137.