ESSAYS ON
THE MICROECONOMICS OF FLOODING: ECONOMIC IMPACTS, BEHAVIOURAL FACTORS AND INDEX–BASED INSURANCE

JOSE ARMANDO
COBIAN ALVAREZ

A THESIS SUBMITTED FOR THE
DEGREE OF DOCTOR OF PHILOSOPHY
AT THE AUSTRALIAN NATIONAL UNIVERSITY

FEBRUARY 2025

Declaration

This thesis is my own work and was written during my Ph.D. studies at the Crawford School of Public Policy, College of Law, Governance and Policy, at The Australian National University. It contains no material previously published or written by another person unless otherwise acknowledged in the text.

The data used in Chapter 2 were generated by the Survey on the Impact of Jakarta Floods, conducted in 2018 and funded by the Australia-Indonesia Centre. The datasets used in Chapter 3 were sourced from the Peruvian National Household Survey (ENAHO) of the National Institute of Statistics and Informatics of Peru (INEI), GEO GPS PERU, the National Aeronautics and Space Administration (NASA), the Environmental Systems Research Institute (ESRI), the Food and Agriculture Organization of the United Nations (FAO), and the Colorado School of Mines. The data used in Chapter 4 were collected through the Household Survey on the Impact of Floods in Lima, carried out in 2023.

The views expressed in this thesis are my own and do not necessarily reflect those of the institutions mentioned above. All errors are my own.

Jose Armando Cobian Alvarez

February 2025

Acknowledgements

This PhD journey has been a profound experience. There were moments when I felt completely overwhelmed and frustrated. Yet, one thing is certain—I was never bored. Each challenge pushed me to my limits, but I persevered, growing in ways that extended beyond academia. Now, despite the struggles, I see the world with a greater sense of wonder and appreciation. And for that, I am truly grateful.

I thank God for allowing me to follow this path. Every step has unfolded as it was meant to.

First and foremost, I would like to express my deepest gratitude to my supervisor, Professor Budy Prasetyo Resosudarmo, for his invaluable guidance and support throughout my PhD. His mentorship has equipped me with the essential tools to refine my research skills and advance my academic development. I would also like to sincerely thank my panel members, Dr. Ryan Edwards and Dr. Firman Witoelar Kartaadipoetra, for their insightful feedback and encouragement, which have enriched my research work.

My sincere gratitude to Professor Blane Lewis, Professor Paul Burke, Professor Ligang Song, and Dr. Arianto Patunru for their constructive comments and input during my presentations at the PhD seminars at the Arndt-Corden Department of Economics. I also thank my discussants for each of my three papers—Ruth Nikijuluw, Dyah Pritadrajati, and Xinzhen Zhang—for their valuable suggestions and remarks.

To my parents, Armando Cobian and Teresa Alvarez, and my siblings, Magaly Cobian and Rafael Cobian, words cannot express how grateful I am for your love, encouragement, and constant belief in me from start to finish. Lastly, I also dedicate this thesis to Ilise Girvan for her patience and support, especially during the difficult moments of my PhD journey. None of my achievements would have been possible without them.

I am also thankful to the National Program of Scholarships and Educational Loans (PRONABEC) for awarding me the Peru Awards Scholarship ("Beca Presidente de la República 2018"); without this generous financial support, I would never have been able to pursue my doctoral degree. In addition, I appreciate the support of the Crawford School student office, particularly Dr. Megan Poore and Tracy McRae, for their academic and non-academic assistance.

I would like to thank the Crawford School of Public Policy and the ANU Tjabal Indigenous Higher Education Centre for the opportunity to work as a tutor for the following courses: Economic Development, Econometric Techniques, Mathematical Methods in Applied Economics, Microeconomics I, Regression Modelling for Actuarial Studies, The Economic Way of Thinking, Economics for Government, Introduction to Spanish I and II, and Continuing Spanish I.

A special thanks to my friends in Australia involved in my journey to achieve my PhD, Haojie Chen, Phillip Drummond, Katia Bonga, Stephanie Courtney-Jones, Luigi Toda, Jennifer Dawes, Isabelle Huang, Zaw Hein, Leonardo Carbajal, Monica Cabrera, Marcos Penteado, Hadrien Brassens, Rosmery Soto, Logea Nao, and a warm thank you to Cristian Gomez for his continuous support and for taking his time to attend and record all my presentations. To my friends in Peru, thank you for looking out for me, Sheyla Montoya, Andrea Ramirez, Cynthia Lopez, Yessica Ochoa, Javier Varillas, Karina Olano, Carolina Lenkey, Lev Requena, Piero Suarez, Erick Cairo, Joel Parreño, Yomeini Perez, Raul Diaz, Rosario Ibarra and Sara Campero.

I also extend my gratitude to my PhD colleagues—Debasish Das, Sulistiyo K. Ardiyono, Kevin Robb, Yuventus Effendi, Abdul Nasri, Inggrid, Riswandi, Christopher Cabuay, Anna Falentina, Agung Widodo, Wawan Cerdikwan, Hali Aprimadya, and Pyan Muchtar—for their camaraderie and encouragement. I am especially grateful to Dr. Sandra Potter for being my ArcGIS Pro tutor during my doctoral research.

I am indebted to all the enumerators and Karen Rocha who helped me conduct the household survey in the communities of Nicolás de Piérola, Sierra Limeña, and San Antonio de Pedregal in the district of Lurigancho–Chosica, Lima; without their assistance, I could not have started writing the fourth chapter of this thesis. I would also like to thank Diego Miranda, Christian Calderon, Vlademir Lozano, and Fred Espinoza from the Research Group Kawsay TIPS for sharing valuable tips and advice during the data collection process.

Abstract

Floods are a global threat, affecting a quarter of the world's population, most of whom live in developing countries. They have severe impacts on welfare and poverty, although coping mechanisms can mitigate these effects. Accurately measuring flood impacts is crucial for assessing their economic consequences and understanding household responses and behaviours. Additionally, substantial losses in urban areas have increased interest in risk transfer products such as index-based insurance, which can offer financial protection, promote sustainable development, and enhance resilience to weather shocks. However, most empirical research on index insurance has focused on rural settings, where farmers are typically insured against income risks from rainfall variation or drought.

This dissertation addresses these gaps by proposing a novel method to estimate the economic impact of floods and by offering evidence on coping mechanisms in a middle-income country such as Peru. It also provides new evidence on factors – such as basis risk, time and risk preferences, price, and trust – that influence the adoption of index-based flood insurance in the urban contexts of Jakarta and Lima.

Chapter 2 analyses demand for hypothetical index-based flood insurance in Jakarta, using household data collected in 2018. Employing a probit specification with plausibly exogenous variables (basis risk, price, and extreme risk aversion), the study finds that demand decreases with basis risk, price, and risk aversion. A policy recommendation highlights the investment in more floodgate stations, particularly in western and southern regions, to reduce basis risk and increase insurance uptake. Chapter 3 examines the effects of the 2017 coastal El Niño floods on welfare and poverty in northern Peru. A novel damage index, constructed using remote sensing within the SWAT model, is combined with five years of panel data in a Difference-in-Differences Event Study. Results show that floods reduced income and expenditure per capita while increasing poverty, especially in urban areas. Households smoothed consumption mainly through disaster relief. The development of disaster insurance that bridges the gap between aid and actual losses –estimated at PEN 818.51 (USD 251.02)– could significantly enhance flood resilience among affected households. Chapter 4 explores the role of economic preferences, price, and trust in shaping demand for index-based insurance in Lima. Using household data from a lab-in-field experiment, I leverage the

exogeneity of time and risk preferences, and price. The study shows that demand rises with time preference (measured by choices of larger, delayed payoffs) but falls with extreme risk aversion and price. Notably, trust in government flood mitigation plans increases willingness to purchase insurance among highly risk-averse households.

Contents

Declaration ... 3
Acknowledgements ... 4
Abstract ... 6
Chapter 1 Introduction ... 14
1.1 Floods and index insurance ... 14
1.2 General objectives ... 16
1.3 Thesis scope ... 17
1.3.1 Chapter 2 ... 17
1.3.2 Chapter 3 ... 18
1.3.3 Chapter 4 ... 18
1.4 Thesis contributions ... 19
1.5 Thesis structure ... 20
References ... 21
Chapter 2 Demand for index-based flood insurance in Jakarta, Indonesia ... 24
2.1 Introduction ... 24
2.2. Theoretical framework and literature review ... 27
2.3 Data ... 33
2.4 Empirical framework ... 40
2.4.1 Identification strategy ... 40
2.4.2 Balance test ... 42
2.4.3 Empirical model ... 48
2.5 Results ... 49
2.5.1 Main results ... 49
2.5.2 Extended analysis ... 57
2.6 Conclusions ... 64
References ... 66
Appendix ... 70
Chapter 3 Under the water: flood impacts and economic dynamics in northern Peru ... 78
3.1 Introduction ... 78
3.2 Related literature ... 81
3.3 The 2017 'coastal' El Niño in Peru ... 83
3.4 Empirical strategy ... 85

3.5 Data 89
3.5.1 Household data 90
3.5.2 Remotely sensed data..........90
3.5.3 Government reports: Identifying flooded districts..........94
3.5.4 Construction of the damage index 94
3.6 Pre-estimation 101
3.6.1 Sample balance analysis 101
3.6.2. Modern event study assumptions analysis..........106
3.7 Results..........110
3.7.1 Main results..........110
3.7.2 Robustness test..........112
3.7.3 Extended analysis 115
3.7.4 Potential channels 117
3.8 Policy recommendation 119
3.9 Conclusions..........120
References..........122
Appendix..........127
Chapter 4 Preferences, trust, and index-based flood insurance: Evidence from Peru .138
4.1 Introduction..........138
4.2 Background..........141
4.3 Data..........144
4.3.1 Lab-in-field survey 144
4.4 Empirical Framework 150
4.4.1 Identification strategy 150
4.4.2 Empirical equation..........156
4.5 Results..........157
4.5.1 Main results..........157
4.5.2 Extended analysis 162
4.5.3 Trust as a mechanism..........167
4.5.4 Heterogeneity analysis..........172
4.6 Policy implications..........175
4.7 Conclusion 176
References..........178
Appendix..........182

Chapter 5 Conclusion 192
5.1 Thesis highlights 192
5.2 General findings 194
5.1.1 Specific findings 194
5.2 Policy insights and recommendations 196
5.3 Limitations 198
5.4 Avenues for future research 199
Survey questionnaires 200
Survey on the impacts of Jakarta floods – 2018 200
Household survey on the impact of floods in Lima – 2023 221

List of tables

Table 2.1 Summary statistics………………………………………………………….. 38
Table 2.2 Flood risk and proximity to floodgate stations……………………………………43
Table 2.3 Flood intensity and flood shocks – distance to floodgate (≤5 km vs. ≥19 km)…...44
Table 2.4 Balance tests……………………………………………………………………45
Table 2.5 Balance tests between flood-free and flood-prone urban villages……………….. 46
Table 2.6 Balance tests - distance to floodgate (≤5 km vs. ≥19 km)………………………...47
Table 2.7 Risk preferences and flood risks………………………………………………… 48
Table 2.8 Index insurance demand among sampled households……………………………50
Table 2.9 Index insurance demand among sampled households using LPM + 1……………52
Table 2.10 Index insurance demand among sampled households excluding covariates…….54
Table 2.11 Distance to hypothetical floodgate stations with flood levels above 2m……….. 56
Table 2.12 Price sensitivity demand and distance to floodgate station in each region……... 61
Table 2.13 Index insurance demand by different subgroups…………………………………..63
Table 2.A.1 Distribution of the degree of risk aversion..71
Table 2.A.2 Index insurance demand among households with real monetary payoffs...........73
Table 2.A.3 Premium rates and demand for index insurance 74
Table 2.A.4 Estimation of Mean WTP (MWTP) 75
Table 3.1 Summary statistics 100
Table 3.2 Balance tests between unaffected and affected areas – before PSM 102
Table 3.3 Exogeneity test: OLS regressions – before PSM 103
Table 3.4 Balance and exogeneity tests – unaffected and affected groups – after PSM.......106
Table 3.5 Double differences between treated and control groups 109
Table 3.6 District of birth and households owing a second house 110
Table 3.7 Impact of flood shocks on household welfare and poverty..................................111
Table 3.8 Impact of flood shocks on household welfare and poverty, Dynamic (TWFE) ...113
Table 3.9 Heterogeneity impacts..116
Table 3.10 Income earners, wage income and disaster relief...118
Table 3.A.1 Exogeneity test of binary switching treatment and index 130
Table 3.A.2 Exogeneity of binary switching treatment and index for "never treated".........131
Table 3.A.3 "Switchers" districts in northern Peru..132
Table 3.A.4 "Always treated" districts in northern Peru..133
Table 3.A.5 "Never treated" districts in northern Peru 133
Table 3.A.6 Impact of floods on household welfare and poverty using 2013 night lights ...135
Table 3.A.7 Impact of floods on household savings 137
Table 4.1 Time preferences..147
Table 4.2 Risk Preferences: Payoffs offered to different investment options.......................148
Table 4.3 Summary statistics 152
Table 4.4 Exogeneity tests: OLS regressions after PSM 154
Table 4.5 Correlation of economic preferences 155
Table 4.6 Balance tests between flood-free and -prone urban communities after PSM 155
Table 4.7 Index insurance demand among surveyed households 160
Table 4.8 Probit specification and OLS model without covariates......................................161
Table 4.9 Effects of levels of patience and expenditure on insurance demand 163
Table 4.10 Factors to determine the success of implementation of index insurance............165
Table 4.11 Trust in institutions and reasons for not purchasing insurance...........................170
Table 4.12 Trust impacts on extreme risk-averse households and insurance uptake............171

Table 4.13 Heterogeneity of the effects of economic preferences..173
Table 4.14 Heterogeneity in the effects of economic preferences on trust...........................174
Table 4.A.1 Distribution of sampled households by choice set sequence 183
Table 4.A.2 Distribution of sampled households by the degree of risk aversion 183
Table 4.A.3 Exogeneity tests: OLS regressions before PSM...186
Table 4.A.4 Balance tests between flood-free and -prone urban communities before PSM.188
Table 4.A.5 Premium rates and demand for index-based flood insurance 190
Table 4.A.6 Cost of implementing index insurance..191

List of figures

Figure 1.1 Global flood hazard frequency and distribution map 15
Figure 2.1 Possible scenarios of households' welfare with index insurance 29
Figure 2.2 Location of study area, sampling sites and reference floodgate stations.............. 35
Figure 2.3 Probability for accepting index insurance by different price levels 40
Figure 2.4 Illustration of spatial proximity 41
Figure 2.5 Price sensitivity of demand and distance to reference floodgate station 57
Figure 2.6 Probability of insurance purchase across levels of risk aversion 58
Figure 2.7 Price sensitivity across levels of risk aversion 60
Figure 2.A.1 Elevation difference between flood-prone and -free areas 70
Figure 2.A.2 Risk aversion difference between hypothetical and monetary games 71
Figure 2.A.3 Extreme risk aversion of demand and distance to floodgate 72
Figure 2.A.4 Outliers regarding the urban villages with flood levels above 2 metres 72
Figure 2.A.5 Village flood occurrences during 2013-2017 76
Figure 2.A.6 Frequency distribution of flood shocks 2013-2017 77
Figure 3.1 SST anomaly in March 2017 along the northern coast of Peru 85
Figure 3.2 Maps of flood damage zones at the district level in northern Peru 99
Figure 3.3 Switcher and non-switcher districts from 2015 to 2019 101
Figure 3.4 Propensity score matching 104
Figure 3.5 Kernel density before and after matching 105
Figure 3.6 Event study plots – de Chaisemartin et al. (2023) 108
Figure 3.7 Event study plots – Dynamic TWFE 114
Figure 3.8 Economic dynamics after the disaster on poverty in urban areas 117
Figure 3.9 Dynamics of income earners, wage income, savings and disaster relief 119
Figure 3.A.1 Timeline of the 2017 coastal El Niño and subsequent flood events 127
Figure 3.A.2 Flowchart of the construction of the flood damage index 128
Figure 3.A.3 Watershed study design for Tumbes river basin 129
Figure 3.A.4 Average covariates over time for treated and control groups 134
Figure 3.A.5 Event study plots using 2013 night lights – de Chaisemartin et al. (2023) 136
Figure 3.A.6 Event study plot on savings (1,0) – de Chaisemartin et al. (2023) 137
Figure 4.1 District of Lurigancho–Chosica vulnerable to flooding 143
Figure 4.2 Location of households and geodynamic barriers, and flood water paths 146
Figure 4.3 Probability of insurance as a function of the non-parametric time preference 162
Figure 4.4 Probability of insurance as a function of the non-parametric risk preference 166
Figure 4.5 Marginal effects extreme and least risk aversion on demand at different prices 167
Figure 4.A.1 Marginal effect of household's WTP for index-based flood insurance 183
Figure 4.A.2 Trust in different types of insurers across urban communities 184
Figure 4.A.3 Trust in gov. development policy and flood mitigation plan 184
Figure 4.A.4 Trust in gov. development policy and mitigation plan across levels of risk ... 185
Figure 4.A.5 Propensity score matching graph 187
Figure 4.A.6 Price sensitivity of insurance demand and different level of patience 189
Figure 4.A.7 Probability of insurance purchase across levels of risk aversion 189

Chapter 1
Introduction

1.1 Floods and index insurance

Flooding is a global threat, with nearly one in four people—around 1.81 billion individuals—at significant risk, 89% of whom live in low- or middle-income countries (Rentschler et al., 2022). The Americas and Asia have been particularly affected by major flood events between 1985 and 2003, as shown in Figure 1.1, where large areas experienced overlapping events (indicated in darker blue). In 2023 alone, floods affected over 32 million people and caused economic losses of US$20.4 billion (CRED, 2023). Climate change is expected to further increase the frequency and intensity of such events (Brunner et al., 2021; IPCC, 2022), exacerbating vulnerability in regions that already bear the brunt of climate-related disasters (Kahn, 2005).

Understanding how floods impact household welfare and poverty, and identifying the mechanisms households and communities use to cope is essential for designing effective policies that protect vulnerable populations and support sustainable disaster resilience strategies. First, empirical studies have linked flood exposure to adverse outcomes such as reduced income, lower consumption, and loss of assets. Most rely on indirect proxies such as rainfall or satellite imagery to measure flood intensity (e.g., Poaponsakorn et al., 2015, Lertamphainont and Sparrow, 2016; Noy et al., 2020; Kocornik-Mina et al., 2020), which often fail to capture the full extent of flood damage due to the complex nature of hydrological processes (Guiteras et al., 2015).

Second, a promising response to flood risk is the use of financial risk transfer tools that provide rapid funding and liquidity after disasters (Kousky, 2019). Among them, index-based insurance (also known as parametric insurance) is a mechanism that relies on predefined weather-related indices associated with losses. These indices are recorded at monitoring stations using historical data. If they cross a predetermined threshold, a payout is triggered. Over the past decade, index-based insurance has gained popularity as a risk management tool, showing significant promise in promoting sustainable development and resilience to weather-related shocks in rural areas of developing countries.

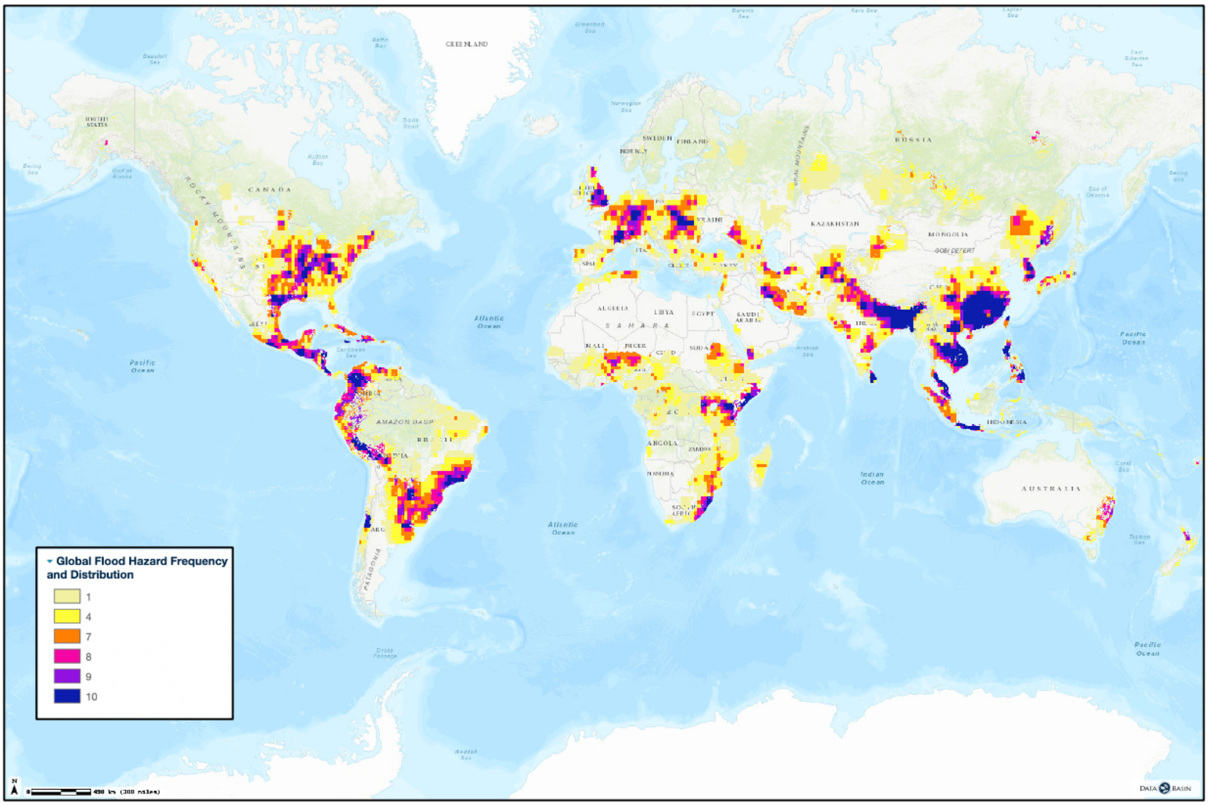


Figure 1.1 Global flood hazard frequency and distribution map

Notes: This flood map is based on a 2.5 × 2.5-minute grid, derived from a global record of extreme flood events between 1985 and 2003, compiled by the Dartmouth Flood Observatory and georeferenced to the nearest degree. The resulting flood frequency grid was classified into 10 categories, each containing approximately the same number of grid cells. Higher grid cell values in the final dataset indicate a greater relative frequency of flood occurrences.
Source: Centre for Hazards and Risk Research (CHRR); Centre for International Earth Science Information Network (CIESIN), Columbia University.

Research on index insurance shows it helps smooth consumption, prevents the sale of productive assets, and enhances investment and credit access among farmers (Kazianga and Udry, 2006; Miranda and Farrin, 2012; Janzen and Carter, 2013; Karlan et al., 2014). For instance, demand for index-based livestock insurance among pastoralists in Mongolia increases in response to adverse winter conditions, indicating that risk exposure during the sales period significantly influences purchasing decisions (Mogge & Kraehnert 2025). Moreover, recent evidence on the long-run effects of this index insurance indicates persistent gains in household production strategies—specifically shifts in herd composition—and in children's education (i.e., human capital accumulation) a decade after initial uptake (Barrett et al., 2024). Theoretical work shows that existing informal risk-sharing mechanisms can affect the demand for insurance (Annan and Datta, 2022; de Janvry et al., 2014; Dercon et al., 2014). Studies exploring the adoption of weather insurance in developing countries have largely focused on factors such as price, product knowledge, trust in the provider, liquidity constraints, financial illiteracy, transaction costs, and risk aversion (Giné et al., 2008; Patt et al., 2009; Gaurav et al., 2011; Cole et al., 2013; Cai et al., 2015; Clarke, 2016; Cai and Song, 2017). However, there remains an urgent need to explore and investigate the features of index-based flood insurance demand for the urban context in developing countries.

This thesis addresses a research gap by developing a new method to assess the economic impact of flood shocks on welfare and poverty, while also offering research evidence on the coping mechanisms used by households in a middle-income country such as Peru. By using hydrological modelling and remote sensing data to construct a damage index and combining it with household survey data, the analysis provides a more granular understanding of the effects of this hazard, especially in urban areas. Moreover, the approach can be replicated for future flood events, with findings that can support central government efforts to respond more quickly, allocate resources where they are most needed, and estimate both the costs and necessary repairs.

On the other hand, the current research provides empirical evidence on the adoption of an index-based insurance product in the urban context of Jakarta, Indonesia, and Lima, Peru. The two studies examine the factors influencing demand for index-based flood insurance in these cities, providing insights into how price, distance to the weather station (a proxy for basis risk), economic preferences (time and risk preferences), and levels of trust influence uptake. Existing research has focused mainly on insurance schemes in agricultural settings, such as rainfall or livestock; however, to the best of my knowledge, there is no empirical evidence on urban contexts—particularly for flood insurance—which present distinct challenges. Urban floods involve different risk exposures and shocks, infrastructure and institutional settings, as well as varied coping mechanisms compared to rural areas, all of which uniquely shape index insurance demand. Consequently, findings from agricultural index insurance cannot be directly applied to the urban context.

1.2 General objectives

This thesis is centred around three main objectives. First, it links empirical analysis to an Expected Utility Theoretical Framework for the demand for insurance, examining the effects of basis risk, premium, and risk aversion on hypothetical index flood insurance take-up rates in Jakarta, Indonesia. Second, it investigates the causal impact of flood shocks on household welfare and poverty in northern Peru, exploring the mechanisms to mitigate the decline in household income and per capita expenditure. Third, it analyses the role of economic preferences and trust in shaping index flood insurance uptake in Lima, Peru. The thesis contributes to the existing empirical literature by evaluating the economic impacts of flooding on household welfare and poverty, and testing whether there is demand for potential countermeasures to mitigate the adverse effects of floods on urban communities in Indonesia and Peru.

1.3 Thesis scope

The thesis focuses on evaluating the impacts of flooding on household well-being and poverty, and studies features of demand for hypothetical index insurance product among households in urban areas of Jakarta, Indonesia, and Lima, Peru.

1.3.1 Chapter 2

This chapter analyses the demand for a hypothetical index-based insurance product designed to cover financial losses and property damage due to floods among households in Jakarta, Indonesia. It focuses on the effects of basis risk – the risk that the payout from the index insurance does not match the actual loss experienced – premium discounts, and risk aversion – the preference for certainty over uncertainty – on insurance uptake. An expected utility framework is used to examine whether basis risk significantly influences the demand for index-based flood insurance. The chapter also explores the impacts of premium discounts and risk aversion on insurance uptake.

The value added by this approach lies in its emphasis on how demand decreases due to basis risk and price. Interestingly and counterintuitively, the results also indicate that higher risk aversion is associated with lower index insurance uptake. This finding is consistent with Clarke (2016), who shows that highly risk-averse individuals are less likely to purchase index-based products due to basis risk and uncertainty around payouts. This model entails testable theoretical predictions that motivate this empirical work.

The study uses household survey data collected across the entire megacity of Jakarta in 2018. The identification strategy is based on the relative exogeneity of basis risk to household characteristics and the plausibly exogenous variables of premium rates and risk aversion. Using a probit specification, the findings indicate that insurance demand among urban households declines as basis risk and premiums increase. In addition, demand decreases at higher levels of risk aversion, demonstrating the reduced appeal of a product with imperfect coverage due to basis risk.

The study uses household survey data collected across the entire megacity of Jakarta in 2018. The identification strategy is based on the plausibly exogenous nature of basis risk, along with plausibly exogenous sources of variation in premium rates and risk aversion. Using a probit specification, the findings indicate that insurance demand among urban households declines as basis risk and premiums increase. In addition, demand decreases at higher levels of risk aversion, demonstrating the reduced appeal of a product with imperfect coverage due to basis risk.

1.3.2 Chapter 3

This chapter assesses the economic impact of the 2017 "coastal" El Niño floods in the northern region of Peru. It proposes a novel damage index as a proxy for the local economic impact, constructed from several remote sensing datasets, including Digital Elevation Models, Land Use and Land Cover, Weather Data, and Night-time Lights. This index is then combined with five years of panel data on socio-economic and dwelling characteristics from the Peruvian National Household Survey (ENAHO). To establish a causal relationship between flooding events and household welfare and poverty status,, a diff-in-diff event study is employed to leverage the exogenous variation in flood events, before and after the coastal El Niño. The results show that households in affected districts experienced a decrease in income and expenditure per capita compared to those in unaffected districts during the period 2015–2019. Additionally, poverty increased as a result of this shock, especially among households in urban areas. Although there was a recovery in income and expenditure per capita in the aftermath of the floods, households smoothed their consumption through the depletion of their savings and the receipt of donations of food and clothing.

1.3.3 Chapter 4

This chapter examines how economic preferences (such as time and risk preferences), price, and trust affect the adoption of a hypothetical index-based insurance product in the megacity of Lima, Peru. Using data from a lab-in-field experiment conducted in the district of Lurigancho–Chosica, I exploit the exogenous variation in time and risk preferences, as well as in price, to estimate the probability of urban households purchasing index-based flood insurance. The experiment includes a repeated discount rate game with hypothetical payoffs to elicit time preferences, a lottery choice to measure risk preferences, and a multiple price list to determine insurance demand at each price. It also includes measures of trust in local institutions and in the insurance product.

This chapter uses an Ordinary Least Squares (OLS) specification that controls for a wide range of covariates, including household-level characteristics, leveraging the detailed survey data. The findings indicate that demand for the insurance product increases with patience but decreases with both price and risk aversion. Consistent with the findings from Jakarta, more risk-averse households are less likely to purchase index insurance. While this result may seem counterintuitive under the Expected Utility Model, further analysis reveals that low trust in traditional insurance providers and in the local government's flood risk management plan discourages adoption, particularly among highly risk-averse households.

1.4 Thesis contributions

This thesis makes meaningful contributions to the empirical literature in each chapter. Chapter 2 introduces two key additions. First, it tests the implications of a theoretical model for index insurance, specifically examining the relationship between insurance uptake and basis risk. It does so by using the distance of a surveyed house to the reference floodgate as a proxy for basis risk. Given that basis risk is under-researched in the developing world (Jensen et al., 2016), the findings provide insight into the demand for a hypothetical insurance product under basis risk conditions in the context of Indonesia. Second, this chapter examines features of demand for index-based flood insurance, including basis risk, premium costs, and risk aversion, within an urban setting (megacity of Jakarta) in a developing country (Indonesia).

Chapter 3 employs a new approach to calculating the local economic damage caused by flood shocks and estimating the impact of this phenomenon on household welfare. Unlike previous studies that exploit variations in rainfall data or use satellite images to identify areas prone to flooding, this chapter constructs a novel damage index that measures the local economic impact using remote sensing data sensitive to surface water. A diff-in-diff event study is then designed to estimate the impacts on income, expenditure per capita, and poverty, as well as coping mechanisms households use to smooth the adverse effects. To date, empirical evidence on the impact of environmental shocks has mostly focused on high- and low-income countries, particularly in rural contexts. Chapter 3 extends the analysis to both urban and rural areas in the context of a middle-income country, such as Peru. Lastly, it examines the heterogeneous effects and economic dynamics of floods on affected districts due to the coastal El Niño floods, using both binary and continuous treatments that switch on and off over time and across groups, with treated units receiving treatment in the same year.

Chapter 4 presents three insights. First, it examines the effects of economic preferences on the demand for a hypothetical index-based flood insurance product, leveraging the exogenous variation in time and risk preferences as well as price, elicited through experimental games. Second, the research incorporates the level of trust in the insurance provider and product. Since previous studies have found limited statistical power in this area, this chapter explores the relationship between trust and economic preferences and their effects on the adoption of index insurance. Finally, the chapter analyses the characteristics of index insurance demand in an urban context, focusing on the megacity of Lima in the developing country of Peru.

1.5 Thesis structure

The remainder of this thesis is as follows: Chapters 2-4 present the three standalone essays on flood shocks impact and index-based insurance to mitigate their adverse effects, and Chapter 5 concludes the thesis. Chapter 2 investigates the features of demand for an index-based flood insurance designed to protect households in Jakarta from the impact of flood events. Chapter 3 studies the economic impacts of devastating floods in northern Peru due to the 'coastal' El Niño in 2017. Chapter 4 explores the role of economic preferences and trust in the development of an index-based flood insurance in the megacity of Lima, Peru. Lastly, Chapter 5 presents highlights of the results and contributions of each paper, then discusses policy implications and future avenues of continuing research.

References

Annan, F, Datta, B 2022, 'Risk, informal institutions, and index insurance', *Journal of Environmental Economics and Management*, vol. 113, pp.1-9.

Brunner, MI, Swain, DL, Wood, RR, Willkofer, F, Done, JM, Gilleland, E, and Ludwig, R 2021, 'An Extremeness Threshold Determines the Regional Response of Floods to Changes in Rainfall Extremes', *Communications Earth & Environment*, vol. 2, no. 1.

Cai, H, Chen, Y, Fang, H, Zhou, L-A 2015, 'The effect of microinsurance on economic activities: evidence from a randomized field experiment', *Review of Economics and Statistics*, vol. 97 no. 2, pp. 287–300.

Cai, J, Song, C 2017, 'Do disaster experience and knowledge affect insurance take-up decisions?', *Journal of Development Economics*, vol. 124, pp. 83–94.

Clarke, DJ, 2016. A theory of rational demand for index insurance. American Economic Journal: Microeconomics, vol. 8, no. 1, pp. 283–306

Cole, S, Giné, X, Tobacman, J, Topalova, P, Townsend, R, Vickery, J 2013, 'Barriers to household risk management: evidence from India', *American Economic Journal: Applied Economics*, vol. 5, no. 1, pp 104–135.

CRED Publication 2023, *Disasters in Numbers - A Significant Year of Disaster Impact*, Centre for Research on the Epidemiology of Disasters, <https://files.emdat.be/reports/2023_EMDAT_report.pdf>.

de Janvry, A, Dequiedt, V, Sadoulet, E 2014, 'The demand for insurance against common shocks', *Journal of Development Economics*, vol. 106, pp. 227-238.

Dercon, S, Hill, RV, Clarke, D, Outes-Leon, I, Taffesse, AS 2014, 'Offering rainfall insurance to informal insurance groups: evidence from a field experiment in Ethiopia', *Journal of Development Economics*, vol. 106, pp. 132-143.

Gaurav, S, Cole, S, Tobacman J 2011, 'Marketing complex financial products in emerging markets: evidence from rainfall insurance in India', *Journal of Marketing Research*, vol. 48, pp. S150–S162.

Giné, X, Townsend, R, Vickery, J 2008, 'Patterns of rainfall insurance participation in rural India' *The World Bank Economic Review*, no. 22, pp.539-566.

Guiteras, R, Jina, A, Mobarak, AM 2015, Satellites, self-reports, and Submersion: exposure to floods in Bangladesh, *American Economic Review: Papers & Proceedings*, vol. 105, no. 5, pp. 232–236.

IPCC, see Intergovernmental Panel on Climate Change Intergovernmental Panel on Climate Change 2022, *Impacts, Adaptation and Vulnerability*, Cambridge University Press.

Janzen, SA, Carter, MR 2013, The impact of microinsurance on asset accumulation and human capital investments: evidence from a drought in Kenya, *ILO Research Paper*, no. 31, International Labour Organization, Geneva.

Jensen, ND, Barrett, CB, Mude, AG 2016, 'Index insurance quality and basis risk: evidence from northern Kenya', *American Journal of Agricultural Economics*, vol. 98, no. 5, pp. 1450-1469.

Kahn, ME 2005, 'The death toll from natural disasters: the role of income, geography, and institutions', *Review of Economics and Statistics*, vol. 87, no. 2, pp. 271–284.

Karlan, D, Osei, R, Osei-Akoto, I, Udry, C 2014, 'Agricultural Decisions after Relaxing Credit and Risk Constraints', *The Quarterly Journal of Economics,* vol. 129, no. 2, pp 597–652.

Kazianga, H, Udry, C 2006, 'Consumption smoothing? livestock, insurance and drought in rural Burkina Faso', *Journal Development Economics*, vol. 79, no. 2, pp. 413–446.

Kocornik-Mina, A, McDermott, TKJ, Michaels, G, Rauch, F 2020, Flooded cities, *American Economic Journal: Applied Economics*, vol. 12, no. 2, pp. 35–66.

Kousky, C 2019, 'The role of natural disaster insurance in recovery and risk reduction', *Annual Review of Resource Economics*, no. 11, pp. 3.1-3.20.

Lertamphainont, S, Sparrow, R 2016, 'The Economic Impacts of Extreme Rainfall Events on Farming Households: Evidence from Thailand', PIER Discussion Papers 45, Puey Ungphakorn Institute for Economic Research.

Miranda, MJ, Farrin, K 2012, ‘Index insurance for developing countries’, *Applied Economic Perspectives and Policy*, vol. 34, no. 3, pp. 391–427.

Mogge, L, Kraehnert, K 2025, ‘Perception and protection: the effect of risk exposure on demand for index insurance in Mongolia’, *Journal of Environmental Economics and Management*, vol. 130, pp. 1-29.

Noy, I, Nguyen, C, Patel, Pooja 2020, Floods and spillovers: households after the 2011 great flood in Thai, *Economic Development and Cultural Change*, vol. 69, no. 2, pp. 829–868.

Patt, A, Peterson, N, Carter, M, Velez, M Hess, U, Suarez P 2009, ‘Making index insurance attractive to farmers’, *Mitigation and Adaptation Strategies for Global Change*, vol. 14, no. 8, pp. 737–753.

Poaponsakorn, N, Meethom, P, Pantakua, K 2015, The Impact of the 2011 floods, and flood management on Thai households, *In Resilience and Recovery in Asian Disasters*, pp. 75–104, Tokio Springer.

Rentschler, J, Salhab, M and Jafino BA, “Flood Exposure and Poverty in 188 Countries,” *Nature Communications*, 2022, 13.

Chapter 2
Demand for index-based flood insurance in Jakarta, Indonesia

2.1 Introduction

Changes in extreme weather and climate events can have large impacts on human health and cause losses to wealth. There has been growing evidence showing that the ongoing trend of climate change contributes to higher temperatures that will exacerbate weather-related events in the coming decades, such as flooding and drought (IPCC, 2012; IPCC et al., 2018; UNDRR, 2019). In turn, this increases the vulnerability of communities to natural hazards, especially those in developing countries characterised by poverty and limited coping capacity. To improve developing countries' capacity, mechanisms such as weather-related insurance are considered to reduce or compensate for economic and financial losses. This risk transfer instrument can speed-up rebuilding and recovery processes by providing post-disaster funding and liquidity soon after the natural event (Kousky, 2019).

Index-based insurance[1] has emerged as a new type of financial risk transfer product helping to overcome issues commonly identified under traditional insurance policies (Barnett and Mahul, 2007; IFAD, 2010). More specifically, it removes the problems of moral hazard (hidden action) and adverse selection (hidden information) when compared to traditional insurance (Barnett et al., 2008; Cole et al., 2013). Two additional advantages are its low transaction costs, due to its reliance on an index—which makes it feasible even when the amounts insured are very small—and the timeliness of payouts, as the insurer only needs to observe the index rather than verify individual claims. This type of insurance has become prominent as a tool to address challenges arising from unsustainable land use and inadequate water resource management, as well as the increasing exposure of populations to flooding. In particular, index-based flood insurance protects against damage resulting from flood incidences due to heavy rainfall. Unlike indemnity-based insurance, where the policyholder receives compensation for verifiable losses, this product pays out claims based on observable and measured flood indices that are correlated with losses suffered by policyholders.

---

[1] This product is classified as a derivative which differs from traditional insurance (known as indemnity-based contracts).

In that context, this study proposes a hypothetical product based on flood-related indices for households located in the megacity of Jakarta, Indonesia, which suffers from annual flooding. The motivation for this proposal lies in the fact that the flood insurance market remains underdeveloped, despite the existence of a legal framework for disaster risk management (Law No. 24/2007). In 2012, the Regional Development Planning Agency (BAPPEDA), in collaboration with international insurers, introduced indemnity-based flood insurance schemes in Semarang, but uptake was limited. Many residents did not perceive insurance as necessary, particularly if they had not personally experienced severe flooding (Soetanto et al., 2020).

More recently, the Government of Indonesia has begun insuring public assets, signalling growing institutional interest in market-based disaster risk financing. However, private uptake remains limited and is largely confined to wealthier individuals and businesses in the Greater Jakarta area. Understanding household-level demand is therefore essential for developing insurance solutions tailored to lower-income and more vulnerable populations. This study contributes to that effort by providing insights for scaling up index-based flood insurance in Jakarta and by identifying key factors—such as basis risk, price, and risk aversion—that influence uptake.

In this study, the index insurance product would compensate urban households against actual losses such as housing maintenance repairs, clean-up costs, income loss, and costs related to evacuation[2]. However, this beneficial aspect of the insurance product also presents a problem: basis risk, the situation whereby the insured faces uncertainty that its actual losses are not fully covered because the index fails to accurately predict the loss, or the policyholder receives a payout when no loss is suffered.

Some experimental studies explore these same factors to assess their collective impact on the demand of smallholder farmers in developing countries purchasing index-based insurance. Their findings show that insurance adoption is negative (and highly sensitive) to prices and distance to the station (basis risk); and demand increases at low levels of risk aversion, then decreases at higher levels (Mobarak and Rosenzweig, 2013; Hill et al., 2013; Cole et al., 2013; Hill et al., 2016).

[2] These are the most common household expenditures due to flooding in Jakarta, based on the individual's responses from the survey.

The purpose of this paper is to test how demand for this insurance product behaves under basis risk, and the response to premium and risk aversion using a hypothetical index-based flood insurance for the urban context of Jakarta, Indonesia. This megacity was selected for the present study due to a number of reasons. First, it is one of the biggest and most populated in the world, and vulnerable from flooding due to the occurrence of annual flood-related disasters (Firman et al., 2011; Cobian and Resosudarmo, 2019). This condition comes as a result of: i) high rainfall intensity; ii) subsidence soil[3]; and iii) inadequate hydraulic infrastructure in the city (Sedlar, 2016). In addition, parts of north Jakarta are sinking at an average rate of 15 centimetres (cm) per year, making Jakarta the world's fastest sinking city (Octavianti and Charles, 2018).

Second, the annual cost of these floods has been significant; for example, the 2007 flood is considered the worst natural shock, as 60% of the city was inundated (World Bank, 2008), 340,000 residents were displaced, and critical infrastructure—such as roads, electricity, and water supply systems—was severely disrupted, causing a total loss of US$565 million (Wijayanti et al., 2017). Third, to improve the capabilities of Jakarta's residents to cope with this annual flood event, there has been discussion on the possibility of developing an index-based flood insurance product. This product may insulate household income and consumption against flood shocks that are exogenous to the household unit.

I achieve the paper's objective by using data from a double-bounded dichotomous choice experiment conducted in 2018, combined with five-year flood data. I use as a proxy of basis risk the distance of a household to the reference floodgate station in the hypothetical insurance product they were offered. The identification strategy exploits this distance as a plausibly exogenous variable to examine its effect on insurance uptake. A potential concern is that floodgates may have been strategically located in areas with higher flood risk, which could influence housing conditions and violate the exogeneity assumption. To address this, I provide both historical and empirical evidence supporting the plausibility of this assumption. Historically, the Dutch administration installed floodgates at locations chosen for engineering purposes as part of Jakarta's drainage system, not based on residential patterns (Sedlar, 2016). Empirically, I find no evidence that floodgates are systematically located in more flood-prone areas or that households closer to them differ in housing conditions, after

3 Jakarta faces significant land subsidence, especially in North Jakarta, due to rapid urbanization and excessive groundwater extraction, increasing flood risk and complicating risk management.

controlling for river proximity. These results support the plausibility of the exogeneity assumption.

In addition, I leverage the plausibly exogenous variation in multiple price list and risk aversion to identify their effects on demand. To the best of my knowledge, it is one of the first studies to examine features of demand for index-based flood insurance in the urban context of a developing country.

The results of this study are in line with those of previous study estimates, showing that demand declines with price and distance to the floodgate station, while insurance uptake decreases only among extremely risk-averse households – a result that may appear counterintuitive under standard Expected Utility Theory but is predicted in the context of index-based insurance, where the presence of basis risk discourages participation (Clarke, 2016). In addition, I find that households located equal to and less than five km away from the reference floodgate station are four times as sensitive to prices (demand is elastic) as those households located equal to or more than 12 km away.

The remainder of the paper is structured as follows. The next section defines index-based insurance, providing the theoretical framework and literature review. It is followed by a section describing the household survey on index-based flood insurance in Jakarta. The next section is on the identification strategy and the econometric framework; then, sections on main results and extended analysis, and lastly the conclusion section.

2.2. Theoretical framework and literature review

Index-based insurance compensates the insured based on pre-agreed weather-related indices – measured with historical information recorded at monitoring stations – that represent actual losses within a geographically defined space (Skees and Barnett, 2006). In agriculture, for example, temperature and precipitation data from local gauges are used to construct indices reflecting drought or heavy rainfall. Compensation is released to the insured when these weather-related indices fall below a predefined threshold, indicating drought, or exceed a higher threshold, indicating excessive rainfall.

With index-based flood insurance, the index is constructed based on data related to the extent and depth of water level, and the amount of time for flooding to subside. Floods near rivers are typically spatially correlated, meaning that rising water levels often cause neighbouring areas along the river to flood simultaneously. Under this approach, all households within a given spatial area holding the same insurance policy receive identical payouts, despite actual losses. Hence, indemnity payments are based on objective, observable, and verifiable variables.

Following the approach of Hill et al., 2016, complemented with elements of Clarke's (2016) work, a standard expected utility model for index-based insurance is adapted to explore predictions on whether insurance demand behaves as expected in response to basis risk, premium and risk aversion. Although this study does not contribute new elements to the theoretical model, it adapts this framework to the context of flooding, which differs in important ways from the agricultural settings typically considered in previous works. The contribution, therefore, lies in the application of theoretical approaches to an underexplored empirical setting such as the urban context of Jakarta, Indonesia.

Consider a representative urban household that is strictly risk-averse[4], with welfare W, and faces two states of the world $S = \{Loss = L,\ Loss = 0\}$; where $Loss = L$ with probability p if the household suffered from a disaster, and $Loss = 0$ with probability $1 - p$, if they did not. In the absence of insurance, the expected welfare is $E[W|S] = p(w - L) + (1 - p)(w - 0) = w - p.L$, and the variance of welfare is $Var[W|S] = E[W^2|S] - (E[W|S])^2 = w^2 - 2pwL + pL^2 - w^2 + 2w(pL) - (pL)^2 = p(1 - p)L^2$. The urban household's preferences over welfare are represented by the indirect utility function $V'(.)$, and they are strictly risk-averse, thus, $V'(.) > 0$ and $V''(.) < 0$.

In the presence of index-based insurance, the maximum payout claimed by an insured household is I when facing losses after a bad natural event occurs. Hence, the expected amount of claim per household is pI. The premium of the index-based insurance is equal to mpI, where m is known as "price multiple" that places it above or below the expected claim. Therefore, when $m = 1$, it is an actuarially-fair price; $m > 1$ is an actuarially-unfair price; and $m < 1$ is a favourable product priced below the actuarially-fair price.

---

[4] A risk-averse household with indirect utility function satisfies Constant Absolute Risk Aversion (CARA). This means as welfare increases a household holds the same amount of money in risky assets.

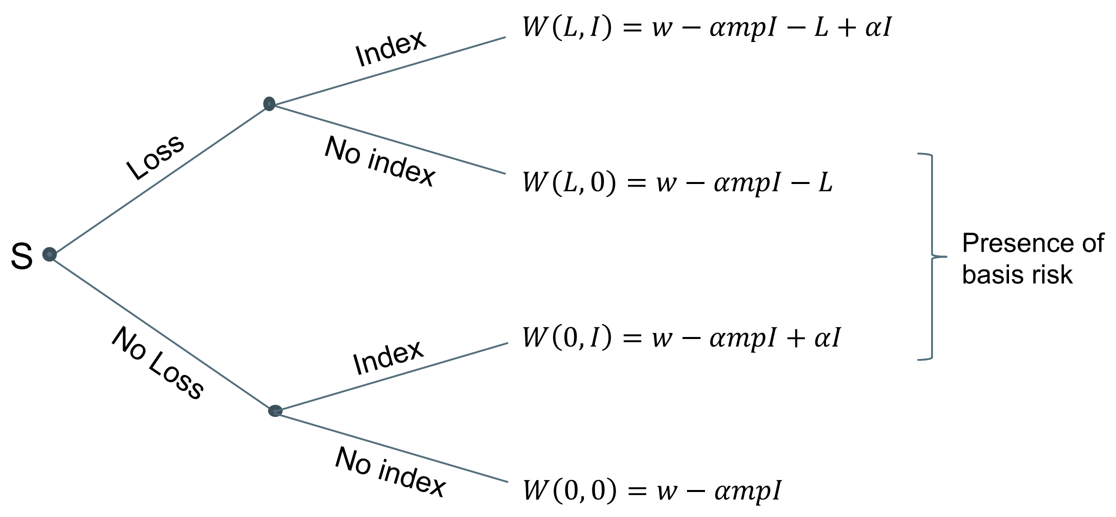


Figure 2.1 Possible scenarios of households' welfare with index insurance when flooding occurs

Notes: Loss = scenario for a household experiencing loss equal to L due to flooding; No Loss = scenario for a household experiencing no loss; Index = scenario for a weather-related index crossing a pre-defined threshold and compensation I released; No Index = scenario for a weather-related index placed below a pre-defined threshold.
Source: Author's illustration.

The urban household can choose an index insurance coverage as high as α, where $0 < \alpha \leq 1$. This urban household would face four possible scenarios when flooding occurs in the neighbourhood area as shown in Figure 2.1 These are combinations of joint events: i) household experiences or not $Loss$ $(= L)$ due to flooding; and ii) whether or not the weather-related index crosses a pre-determined threshold to trigger a payout; i.e., the compensation $(= I)$ is released or not.

The presence of basis risk is associated with a joint probability distribution (r) of some loss experienced by the household due to flooding and the weather-related index placed below a pre-defined threshold – i.e., household facing the worst scenario $W(L, 0)$, or no loss experienced by the household but the weather-related index crosses a pre-defined threshold – i.e., household facing the best scenario $W(0, I)$. In these two scenarios, basis risk is not zero ($r > 0$). In contrast, a perfect index insurance is one that has no basis risk ($r = 0$).
As a result, I have four states of the world (S) and their probabilities $P[W(S)]$:

$P[W(S)] = (Loss = L, Index = 0) = r;$

$P[W(S)] = (Loss = L, Index = I) = p - r;$

$P[W(S)] = (Loss = 0, Index = 0) = 1 - p - r;$

$P[W(S)] = (Loss = 0, Index = I) = r.$

In the presence of insurance, the expected welfare is $E[W|S] = p(w - \alpha mpI - L + \alpha I) + (1 - p)(w - \alpha mpI + \alpha I) = w - \alpha mpI - \alpha I$, and the variance of welfare is $Var[W|S] = p(1 - p)L^2 + 2L(\alpha I)$. Note that if $r = p(1 - p)$, the random variables $Loss$ and $Index$ become identically and independently distributed. This means that the $index\ payout$ is no longer aligned with the actual losses. Therefore, it is required that $r < p(1 - p)$ to ensure that there is at least some dependency between the $Loss$ and $Index$, meaning the $index\ payout$ reflects actual losses to some extent.

Maximisation problem

The household decision problem consists of maximising its expected indirect utility $V(.)$ by controlling α:

$$\max_{\alpha} E[V(W(S))], \tag{1}$$

$$\max_{\alpha} E[W|S] = E[w - \alpha mpI - L(S) + \alpha I(S)] \tag{2}$$

The first-order condition is as follows:

$$E\left[W'(S)\frac{\partial W(S)}{\partial \alpha}\right] = 0 \tag{3}$$

In the absence of basis risk ($r = 0$) and an actuarially-fair price ($m = 1$), the optimal solution is $\alpha^* = 1$. Then, the condition to satisfy the purchase of one unit of insurance is:

$$p(1 - p)LV'(w - \alpha mpI - L + \alpha I) - p(1 - p)LV'(w - \alpha mpI) = 0 \tag{4}$$

Using ($m = 1$) and $\alpha = 1$, I have:

$$p(1 - p)LV'(w - pI) - p(1 - p)LV'(w - pI) = 0 \tag{5}$$

Basis risk

Basis risk (r) indirectly impacts the expected welfare through the optimal choice of α, and increases household welfare dispersion as they face the two extreme states, $W(L, 0)$ and $W(0, L)$. Therefore, purchasing a unit of insurance ($\alpha = 1$) increases the welfare dispersion, leading to the optimal decision of not purchasing insurance at all ($\alpha = 0$). To show this result, I differentiate the first-order condition with respect to α and r as follows:

$$E\left[V''\big(W(S)\big)\left[\frac{\partial W(S)}{\partial \alpha}^2\right]\right]d\alpha \ + \textstyle\sum_S \left(\frac{\partial P[W(S)]}{\partial r} V'\big(W(S)\big)\frac{\partial W(S)}{\partial \alpha}\right) dr \ = 0 \qquad (6)$$

The second term can be expressed as $(1 - mp)L\big[V'(W(0,L) - V'\big(W(L,L)\big)\big] +$ $mpL\big[V'(W(0,0) - V'\big(W(L,0)\big)\big]$. Because the utility function is concave, both bracketed terms are negative, making the entire term negative. Thus, in equation (6), both the first and second terms are negative, which implies that $\frac{d\alpha^*}{dr} < 0$, meaning that insurance demand decreases as basis risk rises.

Premium

The demand for index insurance would also be affected by the magnitude of m (the price factor) when it is different from 1, $m \neq 1$, since the price multiple (m) negatively affects the expected welfare. A household's welfare is negatively impacted by an actuarially-unfair price ($m > 1$) and basis risk ($r > 0$) when the disaster occurs, resulting in a low probability of purchasing the index insurance – basis risk has positive direct impact on welfare variance. Therefore, given possible variations of price and non-price factors – namely the payout relative to the product price, and the degree of basis risk and risk-aversion – in the household's decision problem, the shape of index-based insurance demand becomes an empirical question. I first differentiate the first-order condition with respect to α and m:

$$E\left[V''\big(W(S)\big)\left[\frac{\partial W(S)}{\partial \alpha}^2\right]\right]d\alpha + E\left[V''\big(W(S)\big)(-\alpha pI)\frac{\partial W(S)}{\partial \alpha}\right]dm +$$

$$E\big[V'\big(W(S)\big)(-pI)\big]dm = 0 \qquad (7)$$

Dividing and multiplying the second term by $V'\big(W(S)\big)$ I get:

$$E\left[V''\big(W(S)\big)\left[\frac{\partial W(S)}{\partial \alpha}^2\right]\right]d\alpha \ + \alpha pI\gamma E\left[V'\big(W(S)\big)\left[\frac{\partial W(S)}{\partial \alpha}\right]\right]dm +$$

$$E\big[V'\big(W(S)\big)(-pI)\big]dm = 0 \qquad (8)$$

where γ is the coefficient of risk aversion. The first term is negative; the second term is equal to the first-order condition, multiplied by a constant, and is thus equal to zero; and the third term is negative. Therefore, $\frac{d\alpha^*}{dm} < 0$.

Price sensitivity and basis risk

The responsiveness of demand to price changes depends on the level of basis risk (r). While I do not provide formal proof of this relationship, the reasoning is that demand is more elastic when basis risk is low than when it is high. Specifically, demand elasticity tends to increase as basis risk approaches zero. Conversely, when basis risk is extremely high, demand for index insurance declines to near zero and becomes unresponsive, particularly when the insurance product is priced at or below the actuarially fair price ($m \leq 1$).

Risk aversion

The study undertaken by Clarke (2016) shows that in the presence of basis risk ($r > 0$) with an actuarially-fair price insurance product ($m \leq 1$), it is expected to see a downward demand curve with respect to levels of household risk aversion (from least to extreme risk-aversion). In any case, I predict a least risk averse household willing to purchase insurance ($\alpha = 1$), while an extremely risk-averse household unwilling to purchase insurance ($\alpha = 0$). The former does not expect the worst-case scenario to occur $W(L = Loss, Index = 0)$, where welfare decreases with index insurance (α). In contrast, the latter assumes the worst-scenario will occur and would be unwilling to buy insurance, as they are not prepared to accept a welfare reduction under such circumstances $W(L = Loss, Index = 0)$.

There exists experimental research focusing on the impact that price and non-price factors have on index-based insurance demand, including those in developing countries (Giné et al., 2008). Index-based insurance literature, focusing on analysis of paying a price in return for a future payout using discounts randomly allocated among households to generate exogenous price variation for insurance uptake, shows that, as prices decline via incentives, the probability of purchase increases. This is exacerbated when index-based insurance is below or at an actuarially-fair price, showing that demand is clearly sensitive to these variations for risk-averse households (Cole et al., 2013; McIntosh et al., 2013; Takahashi et al., 2016; Tadesse et al., 2017; Jensen et al., 2017). By contrast, for insurance products with an actuarially-unfair price – with or without discounts, household demand increases with risk-seeking and decreases with risk-averse households (Hill et al., 2016; Clarke, 2016).

Basis risk, however, has been under-researched within index-based insurance literature in the developing country context (Jensen et al., 2016), despite being a major issue substantially reducing demand for insurance. Few studies posit approaches to proxy for basis risk; for example, Giné et al. (2008) use *accumulated rainfall* to consider when a farmer decides to sow seeds[5], and *share of cultivated land used for castor and groundnut crops*[6] to infer if the farmers' planting decision coincides with precipitation measurements collected at the gauge. Another proxy – used in Deng et al. (2007), Gaurav and Chaudhary (2020), and Ceballos and Robles (2020) – is the *degree of correlation* between two sets of weather indices collected at the weather station and the household's plot, finding that basis risk decreases when the coefficient is closer to 1. Finally, Mobarak and Rosenzweig (2013), McIntosh et al. (2013), Hill et al. (2016), and Sibiko et al. (2018) use *distance to the weather station* to measure the basis risk degree of index-based insurance, finding that pay-outs for shorter distances are more closely correlated with actual losses of the insured. A similar proxy is used in this study – distance of a house to the reference floodgate station.

2.3 Data

Field survey

A household survey was conducted for a month between February and March in 2018 to collect basic information on household characteristics, including age, gender, number of members, years of education, income and consumption, and housing ownership; geographic characteristics; and experience with flood shocks. The survey also included a specialised module regarding demand for a hypothetical index-based flood insurance product, premium rates, discounts, and risk attitudes.

Given that areas in Jakarta are exposed to different flood risks[7], I classified each *Rukun Warga – RW* (urban community), or *kelurahan* (neighbourhood), in Jakarta into three categories: i) almost never experienced flood[8]; ii) occasionally experienced flood; and iii) always experienced flood. This classification used data on flood water level at RW level collected by the *Badan Penanggulangan Bencana Daerah* (BPBD or Jakarta Regional

---

5 This proxy is a dummy equal to one if a farmer decides to sow based on accumulated rainfall (due to the approximate 2-month verification delay following rainfall occurring and the government making a payout), and zero if instead the decision is based on other factors such as soil moisture and advice from other farmers (which are non-reliant on government verification).

6 If a household grows castor or groundnut crops, which are related to the presence of low rainfall and associated with the insurance design product, they were more likely to purchase insurance.

7 For houses that do experience flooding, the flood level often reflects how high or low the house is relative to its surroundings. Houses at lower elevation tend to have higher flood water levels, and vice versa.

8 These houses might be at very different elevation levels. Some could be located well above flood-prone areas (e.g., built on a hill or far from water sources), making them unlikely to flood even during extreme events. Others might be just barely above the flood line, narrowly avoiding flooding during a specific event. This variability in elevation among non-flooded houses can make it challenging to compare them directly with flooded houses. However, the data show that the elevation distribution of flood-prone and flood-free houses overlap significantly, indicating that both groups share similar elevation ranges (see Figure 2.A.1 in the Appendix). To account for these differences, I include elevation as a variable in the specification.

Disaster Management Agency)[9] for the period 2013-2016. In collaboration with the Indonesian statistical agency, Statistics Indonesia, 1,200 households were sampled, equally distributed across "almost never flooded", "occasionally flooded" and "always flooded" areas. A total of 836 households completed the survey, of which 258 are in the "almost never flooded" category, 289 in the "occasionally flooded" category, and 289 in the "always flooded" category. Figure 2.2 maps the household locations of survey participants in Jakarta, and the areas covered including within and outside of flood hazard zones, and closer to and far from main rivers.

A unique hypothetical index-based flood insurance product was offered in each surveyed household. The product would pay a fixed amount (IDR 10 million[10]) to cover losses caused by a flood event associated with the river water height crossing a certain threshold (predetermined trigger)[11] corresponding to a particular floodgate station[12]. The hypothetical product was designed in a way that payouts and perils are covered in the area related to the reference floodgate station, but with each station having a different predetermined trigger (flood index levels associated with the river water level). Households were free to express their willingness to pay (or not) for the hypothetical index-based flood insurance product.

The flood indices associated with the river water level relied on 14 reference floodgate stations distributed across the five regions of Jakarta: eight are located within the megacity, and six are at territory borders (see map in Figure 2.2)[13]. However, six floodgate stations were not included in this study as a result of simple random sampling when conducting the survey. No household surrounding these six floodgate stations were chosen to be surveyed.

---

9 A flood event is defined as an overflow of river water with a height of 40 cm and above. According to Cobian and Resosudarmo (2019), approximately from this flood height, households start experiencing property damage associated with flooding. In addition, I did not use data on elevation with respect to sea level to allocate urban villages to any category because of the significant variation across locations with houses that have or have not experienced floods in the 2013-2017 period.

10 It represents the average maximum amount spent to cover household losses such as housing repairs, clean-up costs, and income loss when flooding occurs. This information was obtained from focus group discussions involving people who live in some flood-prone urban communities in Jakarta.

11 The thresholds at which payouts are triggered vary from 200 to 950 cm.

12 Each floodgate station has a particular river water level that was determined by the observed measurement during the 2007 flood event in Jakarta. This is known as the critical level of *Siaga 1* (highest alert).

13 Flood indices consider water level information from fluvial, pluvial and coastal flooding in Jakarta. Floodgate stations record river water levels that are identified with flood heights measured in neighbourhoods affected by any type of this natural hazard.

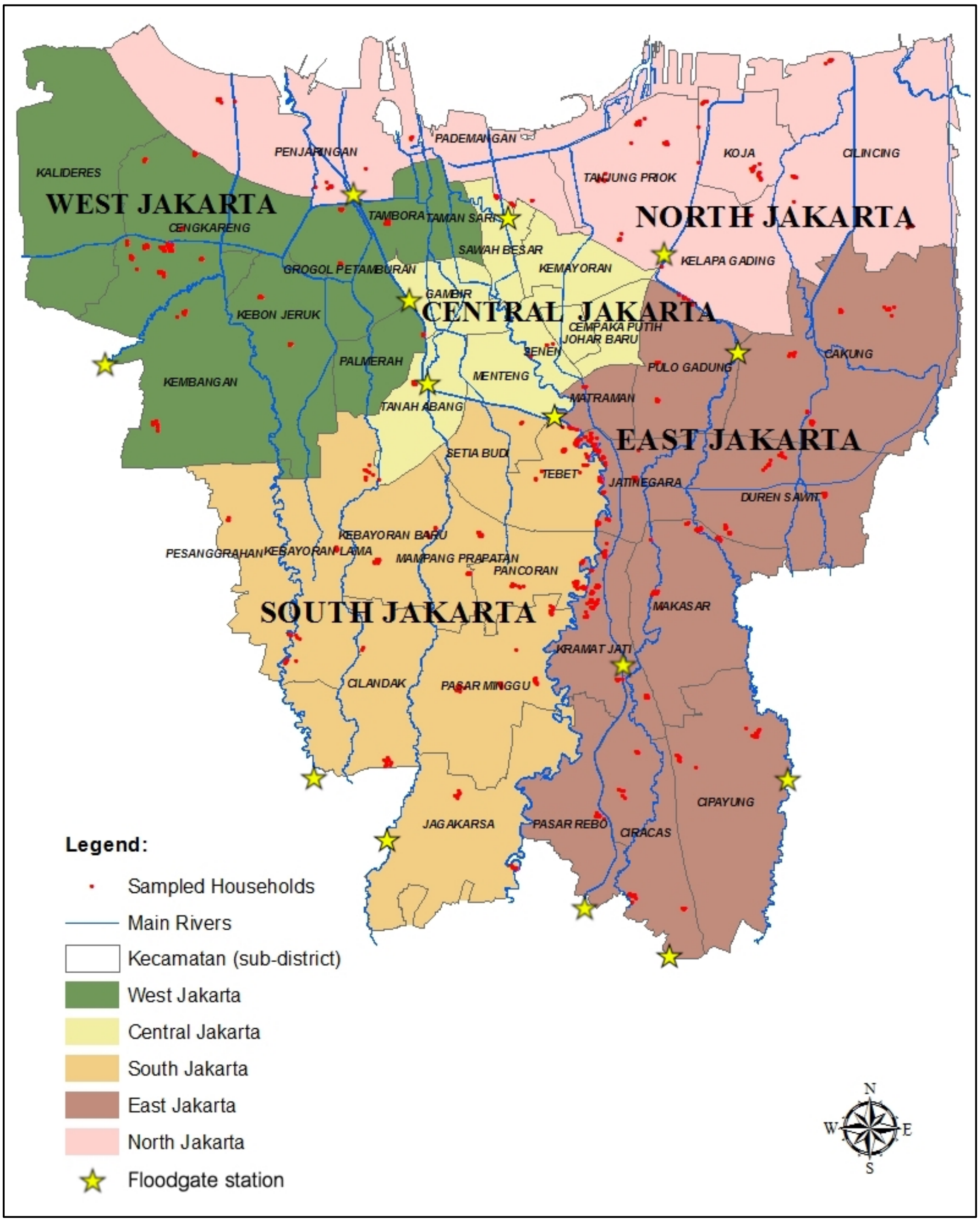


Figure 2.2 Location of study area, sampling sites and reference floodgate stations

Notes: This figure shows the regions of the study area, the households who participated in the survey, and the location of the reference floodgate stations used for triggering payouts of the hypothetical index-based flood insurance.
Source: Author's illustration.

The degree of basis risk associated with the hypothetical index-based flood insurance is measured by using distance of a house to the reference floodgate station (as a proxy). The survey dataset provides the geographic location of households in the study area (x-and-y coordinates expressed as longitude and latitude)[14], whereas locations for existing floodgate stations used as reference stations to trigger payouts were obtained freely from the *Posko Banjir Online* (Flood Information Center Online) website[15].

Surveyed households were referenced to a closer floodgate station if they were downstream from that floodgate station in a particular river; or to a unique, nearest floodgate station in a particular area.

Additionally, surveyed households were offered randomly varying monthly premiums. First, they were randomly offered 1 of 5 different premiums[16] to measure their willingness to purchase. Then, if a household did not choose to purchase the index-based insurance product, it received a price discount of 50 percent; otherwise, it faced a price increase of 50 percent. This method follows the double-bounded dichotomous choice contingent valuation (Hanemann et al., 1991) that improves efficiency over discrete choice models. This hypothetical insurance product, however, was priced at an actuarially-fair price of IDR 167,000[17].

A Binswanger-style lottery choice model (Binswanger, 1980) was utilised in the experiment to measure individual risk preferences and understand how each respondent makes decisions under uncertainty[18]. More recent methods, such as the Eckel-Grossman approach (2002), offer similar simplicity while addressing limitations of the Binswanger design. It presents six lotteries with equal expected values and 50/50 outcomes, requiring a single choice. This

---

[14] There are 81 pairs of missing coordinates in the data due to problems with GPS signal during the survey. Information on household location at the *Rukun Warga* level (Urban Community Unit) and the x-y coordinates of the centroid of these urban community units obtained from OpenStreetMap website are available for this study. By using these sources of information, I fill in the missing values on coordinates.

[15] The information about geographical location of each floodgate station used in this study is available at: http://poskobanjirdsda.jakarta.go.id/map_fullscreen.aspx.

[16] The randomised monthly prices chosen were IDR 10,000 (USD 0.69), IDR 50,000 (USD 3.45), IDR 100,000 (USD 6.91), IDR 200,000 (USD 13.81) and IDR 500,000 (USD 34.53). The base price was obtained from focus group discussions involving people from some flood-prone urban communities in Jakarta before conducting the survey. These urban communities have the "Urban Community Flood Savings" which consists of collecting IDR 10,000 monthly per household in a particular urban community and uses the saved amount to help affected households during flooding.

[17] The actuarially fair price is the price where the insurance premium exactly matches the expected payout. It is calculated using the following formula: $Premium = (Loss * Probability)$, where $Loss =$ IDR 10 million and $Probability = 20\%$, as flood disasters occurred in the years 2002, 2007, 2013, and 2017.

[18] This model included two rounds, hypothetical payments and real payments. Under the game, each respondent chooses one of six investment options. These include IDR 20,000 (USD 1.38) – IDR 20,000; IDR 16,000 (USD 1.10) – IDR 25,000 (USD 1.73); IDR 12,000 (USD 0.83) – IDR 30,000 (USD 2.07); IDR 8,000 (USD 0.55) – IDR 35,000 (USD 3.42); IDR 4,000 (USD 0.28) – IDR 40,000 (USD 2.76); and IDR 0 – IDR 50,000 (USD 3.45). Then, a 'which-hand-is-it in' type game is played with two marbles (blue and yellow), with a probability of 50 percent to yield a return. The respondents are asked to choose a hand. Finally, respondents receive the payment associated with the blue or yellow marble.

allows for cleaner inference on risk preferences and classifies individuals into a few risk types.

While both methods use binary lotteries that vary in risk, key differences include Binswanger's inclusion of dominated options, a nonlinear relationship between risk and return, and variation in expected values. The decision to use the Binswanger approach in this study was motivated by its widespread application in similar empirical settings and its suitability for the study population. Specifically, it has been employed in fieldwork in developing countries due to its intuitive framing and suitability for rural and low-literacy populations (e.g., Bauer, Chytilová, and Morduch 2012; Chowdhury, Sutter, and Zimmermann 2022).

Figure 2.A.2 and Table 2.A.1 in the Appendix report that the tendency of participants to move to a riskier option (increase in 27% compared to the reduction of 6.3% for more risk averse) could have been driven by the fact they were trying to win as much cash as possible, rather than revealing a true preference for risk. Therefore, in this study, I present estimates for the relationship between risk aversion (measured through a hypothetical lottery survey) and demand for index-based flood insurance. This is explained by two factors. Firstly, risk preference measures under hypothetical payoffs are consistent with the expected utility theory. Secondly, I do not see significant differences in estimates when I replicate these results using monetary payoffs.

Table 2.1 presents the summary statistics of the households in the survey and the weather-related variables for the analysis. The categorisation of flood risk – "almost never", "occasionally", and "always" flooded – is reflected in the variables of flood shocks and flood-related expenses in 2017, supporting their validity as proxies for hazard exposure. Additionally, households in the "always flooded" areas are, on average, furthest from the floodgate station but closer to the nearest river. In Section 2.4.1, I examine whether the index (signal) predicts flood hazard based on the distance between the household and the floodgate. This distance-based index underpins the design of the hypothetical index-based insurance product used in the experiment, where basis risk is a key concern.

Table 2.1 Summary statistics

	(1) Almost never flooded areas (258 obs.)		(2) Occasionally flooded areas (289 obs.)		(3) Always flooded areas (289 obs.)		(4) Total (836 obs.)	
	mean	sd	mean	sd	mean	sd	mean	sd
Panel A: Geographic characteristics								
Distance to closest floodgate station (km)	8.524	6.468	7.830	4.313	12.274	3.590	9.580	5.259
Distance to nearest river (km)	0.788	0.679	0.791	0.889	0.192	0.179	0.583	0.711
Elevation (metres above sea level)	16.194	13.746	19.691	18.363	17.409	8.573	17.823	14.209
Panel B: Premium rates								
First question (1,0)	153.411	145.901	157.993	156.787	146.747	148.026	152.691	150.367
Follow-up question (1,0)	91.647	95.371	94.464	84.086	91.851	97.363	92.691	92.246
Panel C: Household head risk preference								
Extreme risk aversion (hypo. Lottery)	0.318	0.467	0.291	0.455	0.311	0.464	0.306	0.461
Extreme risk aversion (real lottery)	0.291	0.455	0.294	0.456	0.277	0.448	0.287	0.453
Panel D: Household characteristics								
Electricity (1,0)	0.996	0.062	0.997	0.059	0.997	0.059	0.996	0.060
Asbestos roof (1,0)	0.450	0.498	0.464	0.500	0.498	0.501	0.471	0.499
Storey (number)	1.256	0.446	1.219	0.422	1.394	0.510	1.291	0.467
House ownership (1,0)	0.764	0.426	0.772	0.421	0.799	0.401	0.779	0.415
Female (1,0)	0.209	0.408	0.183	0.388	0.183	0.388	0.191	0.394
Age (years)	49.651	10.886	50.851	12.074	51.035	12.105	50.544	11.733
Married (1,0)	0.806	0.396	0.813	0.390	0.803	0.399	0.807	0.395
Education (years)	10.318	3.051	9.471	3.480	9.603	3.209	9.778	3.275
Household members (number)	3.702	1.392	3.709	1.431	3.668	1.605	3.693	1.480
Household expenditure (million IDR)	3.769	2.468	3.468	1.860	3.121	1.581	3.441	2.000
Public health insurance (1,0)	0.826	0.380	0.841	0.366	0.830	0.376	0.833	0.374

Panel E: Weather and flood experience								
Rainfall in 2015 (mm)	1,705.50	126.23	1,715.22	160.04	1,677.68	109.84	1,699.25	134.76
Flood expenses in 2017 (million IDR)	0.272	1.368	0.211	0.363	0.424	1.084	0.303	1.017
Flood shocks (number)	4.109	3.164	4.536	2.918	8.581	2.183	5.803	3.432
Panel F: Perception of flood								
Flood risk (1,0)	0.756	0.430	0.803	0.399	0.782	0.414	0.781	0.414
Flood cycle – every 5 years (1,0)	0.581	0.494	0.599	0.491	0.439	0.497	0.538	0.499
Panel G: Mitigation strategy								
Mitigation (1,0)	0.326	0.470	0.304	0.461	0.356	0.480	0.329	0.470

Note: Elevation, precipitation and flood shocks are reported at the neighbourhood level (*kelurahan*).
Source: The Impact of Jakarta Floods Survey (2018).

Figure 2.3 shows the probability of accepting to buy index-based flood insurance at different price levels. As the monthly premium goes up, the probability of insurance purchase declines.

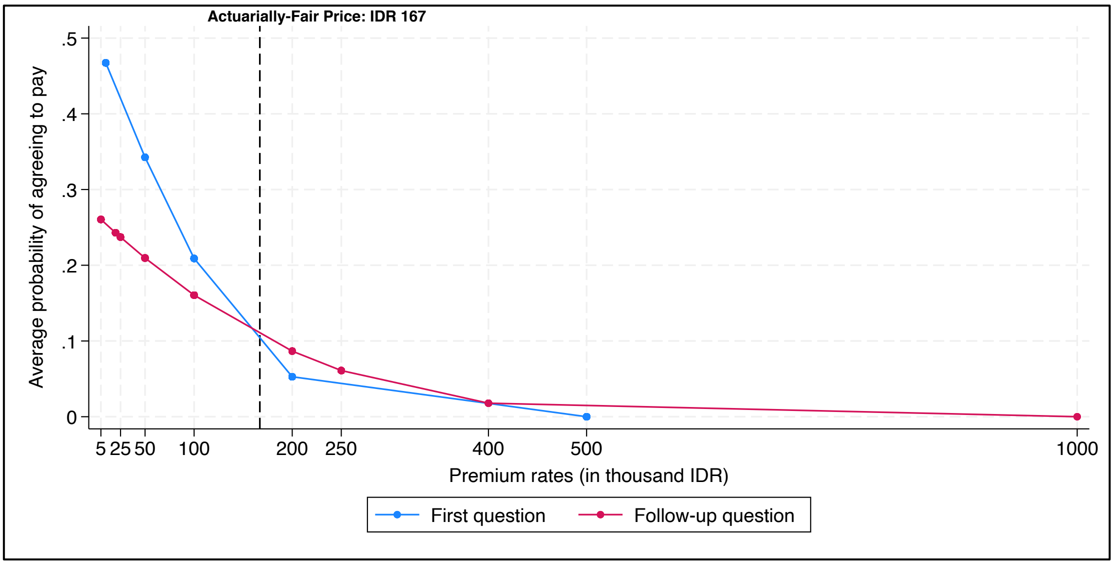


Figure 2.3 Probability for accepting index-based flood insurance by different price levels

Notes: This figure shows the probability of insurance purchase for the first question and the follow-up question under a dichotomous choice method.
Source: The Impact of Jakarta Floods Survey (2018).

2.4 Empirical framework

2.4.1 Identification strategy

This study takes inspiration from Hill et al. (2016) who used distance to the reference station, premium rates, and household risk aversion in a field experiment to explore the predictions of an expected utility model on insurance demand. As a proxy for basis risk, I use the distance of a house to the reference floodgate station and assume it is plausibly exogenous variable after I control for distance of a house to the nearest river. Households' decision on where to live either far from or closer to the river, is mainly driven by income and amenities, such as affordable housing and adequate basic services[19]. By contrast, distance to the floodgate station is typically ignored.

Figure 2.4 illustrates the spatial proximities, that is, the distance (d_i) of house i (H_i) to the reference floodgate station (F_s), and the distance to the nearest river (d_r) – which represent the shortest straight-line (Euclidean) distance from the household's location to the nearest river. The former is the perpendicular distance to the floodgate station and differences in x-

[19] For example, low-income families are often located in densely populated areas that are more likely exposed to flood risks (Jha et al., 2012). In the case of Jakarta, many of the poor have developed large informal settlements along the waterways and rivers (Texier, 2008).

and-y coordinates (a_i and b_i, respectively) to the floodgate station. The latter is the straight-line distance between each sampled household and the nearest river.

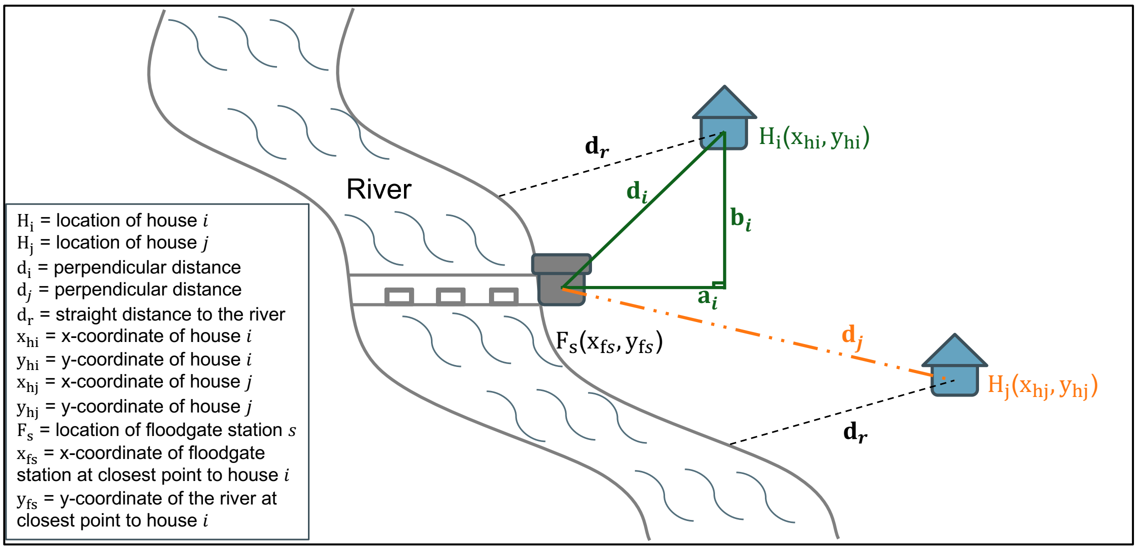


Figure 2.4 Illustration of spatial proximity

Source: Author's own illustration.

In addition, in Figure 2.4, the distance d_i of house H_i to the reference floodgate station F_s is shorter than the distance d_j of house H_j to the same floodgate station F_s where[20] flood indices are measured for both houses, despite the fact that H_j would potentially experience different flood conditions compared to H_i. This creates basis risk where H_j may not receive compensation even though they experience losses from flooding, simply due to the fact that the floodgate station records a different level of flooding (e.g., lower). Therefore, the key identifying assumption is that index-based flood insurance demand is negatively affected by the distance to the reference floodgate station, given that distance to the river is controlled in the estimated model.

In relation to premium rates, one of five price levels were randomly presented to each respondent for the hypothetical index-based flood insurance product. In theory, each sampled individual had an equal probability of being assigned any of these prices. Then, exogenous variation in the price was introduced by randomly allocating a price discount or price increase across urban households, under a double-bounded dichotomous choice. I test this exogeneity by performing balance tests on a number of covariates. Therefore, the

[20] This means that when a house is closer to a floodgate station, the policy is aligned to the flood information nearby (flood height and damage); whereas when a house is further from a floodgate station, the policy is still aligned to the floodgate station information but not local conditions.

identification strategy used in the study assumes the premium rates have no statistically significant effect on these covariates.

Referring to risk aversion, index-based insurance literature has treated risk preferences as exogenous using Binswanger's (1980) lottery choices[21] in field experiments to infer the levels of risk. Recent empirical research has found that risk aversion may be influenced by natural shocks, including flooding (Chuang and Schechter, 2015; Schildberg-Hörisch, 2018); however, in the experiment, using data of household extreme risk aversion and flood shocks, I find no statistically significant relationship[22].

In the estimated model, furthermore, I include region and flood intensity fixed effects–where flood intensity is based on the severity of flood events, categorised as "almost never," "occasionally," and "always" flooded areas–, so that the identifying variations do not disproportionately rely only on flood-prone urban villages where the impact of flooding is larger. Finally, I account for the variation of flood levels between 2013 and 2017 within a particular urban village area by calculating the monthly flood shocks[23]. This controls any weather-related influence such as floods on the household decision to purchase index-based flood insurance.

2.4.2 Exogeneity and balance tests

As noted above, distance to reference weather stations has been used to proxy basis risk in index insurance, under the assumption that these stations were randomly-located. In the urban context of Jakarta, I provide some evidence that supports this assumption. Historically, between the late 1800s and 1945, the Dutch administration installed floodgate stations at randomly selected locations as part of the drainage system within the Ciliwung river basin and main drains, to control both the flow of water from upstream and the volume of water entering Batavia (Sedlar, 2016) – the former colonial port from which Jakarta has expanded. Empirically, I show that the number of flood shocks (as a proxy for flood risk) is lower near

---

21 Each respondent chooses one lottery (investment option) out of six, and a 'which-hand-is-it in' type game with two marbles (blue and yellow) is played. Then, respondents are asked to pick the hand with the blue marble or with the yellow marble (each with a probability of 50%), and the respondent receives the payment accordingly. The investment options are: 1) IDR 20,000 (blue)-IDR 20,000 (Yellow); 2) IDR 16,000-IDR 25,000; 3) IDR 12,000-IDR 30,000; 4) IDR 8,000-IDR 35,000; 5) IDR 4,000-IDR 40,000; and 6) IDR 0-IDR 50,000.

22 The OLS estimate of the effect of flood shocks on extreme risk aversion is -0.117, with a p-value of 0.595, after controlling for the distance from the surveyed household to the nearest river and region fixed effects.

23 Shocks are relative to the subdistrict level long-run flood in a particular month (they are defined at a monthly definition). The estimation of monthly flood shocks follows the approach of Karadja and Prawitz (2017) (see Appendix Section B for more details about the calculation).

floodgate stations, as shown in Table 2.2. In fact, households located further from floodgates, on average, experience more flood shocks and tend to be located in the "always" flooded areas. These findings support the claim that floodgates were placed at random locations, most likely determined by the engineering needs of the Dutch colonial city.

Table 2.2 Flood risk and proximity to floodgate stations

Model: OLS	Dependent variable: Flood shocks					
	(1)		(2)		(3)	
Distance to floodgate station (km)	0.274	***			0.196	***
	(0.020)				(0.020)	
Flood intensity:						
- Almost never flooded areas (1,0)			-0.058		-0.166	
			(0.229)		(0.217)	
- Always flooded areas (1,0)			3.331	***	2.485	***
			(0.245)		(0.247)	
Distance to nearest river (in logs)	Yes		Yes		Yes	
Region FE	Yes		Yes		Yes	
Number of observations	836		836		836	

Notes: This table presents OLS regression estimations of flood shocks (proxy for flood risk) on distance to the reference floodgate station and flood intensity regarding "almost never" flooded and "always" flooded areas. All specifications include distance to nearest river and region fixed effects. Asterisk (*), double asterisk (**), and triple asterisk (***) denote variables significant at 10 percent, 5 percent, and 1 percent, respectively.
Source: The Impact of Jakarta Floods Survey (2018).

To ensure the effectiveness of index insurance, it is important to assess whether the index accurately signals flood risk, at least in the proximity of the floodgate stations, since there is some degree of basis risk in the product. I test this by regressing reported flood shocks on flood intensity categories. In Table 2.3, among households located within 5 km of a floodgate station, those in "always" flooded areas report significantly more flood events than those in "almost never" flooded areas. In contrast, among households located 19 km or more from a station, neither category predicts flood shocks. These results confirm that the index captures variation in flood risk exposure near stations, but its predictive power declines with distance – consistent with the presence of basis risk.

Table 2.3 Flood intensity (signal) and flood shocks – distance to floodgate (close<=5 km vs. far>=19 km)

Model: OLS	Dependent variable: Flood shocks			
	(1) ≤ 5 km from floodgate		(2) ≥ 19 km from floodgate	
Flood intensity				
- Almost never flooded areas (1,0)	0.883	**	0.410	
	(0.418)		(1.420)	
- Always flooded areas (1,0)	7.950	***	-0.025	
	(2.501)		(0.706)	
Distance to nearest river (in logs)	Yes		Yes	
Region FE	Yes		Yes	
Number of observations	170		60	

Notes: This table presents OLS regression estimations of flood shocks (proxy for flood risk) on distance to the reference floodgate station and flood intensity regarding "almost never" flooded and "always" flooded areas. All specifications include distance to nearest river and region fixed effects. Asterisk (*), double asterisk (**), and triple asterisk (***) denote variables significant at 10 percent, 5 percent, and 1 percent, respectively.
Source: The Impact of Jakarta Floods Survey (2018).

On the other hand, the identification strategy relies on the exogeneity of distance to the reference floodgate station, premium rates, and extreme risk aversion. I test this assumption by performing balance tests of these three variables on a number of observable variables at the household level. Table 2.4 displays the outcome of these tests. Most covariates are statistically insignificant; although, by random chance, I should expect some variables to be correlated with the study variables (i.e., female and household expenditure). Therefore, I conclude that distance to the reference floodgate station, premium and extreme risk aversion variables are plausibly exogenous.

Table 2.4 Balance tests

Sample: Dependent variable	Distance to floodgate (km)			Premium rates (thousand IDR)		Extreme aversion (1,0)		
Electricity (1,0)	0.0004	(0.0005)		0.0000	(0.0000)	0.0054	(0.0046)	
Asbestos roof (1,0)	0.0034	(0.0038)		-0.0001	(0.0001)	-0.0215	(0.0378)	
Storey (number)	0.0012	(0.0035)		-0.0001	(0.0001)	-0.0363	(0.0346)	
House ownership (1,0)	-0.0001	(0.0032)		0.0000	(0.0001)	-0.0311	(0.0316)	
Female (1,0)	-0.0089	(0.0030)	***	0.0000	(0.0001)	-0.0547	(0.0299)	*
Age (years)	0.1140	(0.0890)		-0.0029	(0.0027)	-0.8489	(0.8914)	
Married (1,0)	0.0032	(0.0030)		0.0000	(0.0001)	0.0186	(0.0300)	
Education (years)	-0.0234	(0.0248)		-0.0005	(0.0008)	0.0640	(0.2488)	
Household size (number)	-0.0125	(0.0111)		0.0003	(0.0003)	0.0776	(0.1112)	
Household expenditure (million IDR)	-0.0576	(0.0147)	***	-0.0006	(0.0005)	-0.0957	(0.1482)	
Public health insurance (1,0)	-0.0042	(0.0028)		0.0000	(0.0001)	0.0191	(0.0283)	
Flood expenditure in 2017 (million IDR)	0.0071	(0.0077)		0.0001	(0.0002)	-0.0752	(0.0768)	

Notes: OLS regressions. Each result represents a separate regression of the dependent variable on distance to the reference floodgate station, premium rates, and extreme aversion (which are the study variables). Each regression includes distance from the surveyed household to nearest river and region fixed effects. Asterisk (*), double asterisk (**), and triple asterisk (***) denote variables significant at 10 percent, 5 percent, and 1 percent, respectively.
Source: The Impact of Jakarta Floods Survey (2018).

In addition, I conduct balance tests between flood-free and flood-prone urban villages[24]. Households located in a flood-prone area exhibit few significant differences compared to other households outside the flood zone, as shown in Table 2.5. Average differences are generally small: out of 12 variables spanning demographic information, only 2 exhibit significant differences at the 1% level. I thus proceed to estimate the effects of basis risk, premium and extreme risk-aversion on uptake using equation (9) in the next section.

Table 2.5 Balance tests between flood-free and flood-prone urban villages

	Flood-free group	Flood-prone group	Single difference	N	p-value	
	(1)	(2)	(3)	(4)	(5)	
Electricity (1,0)	0.996	0.997	0.000	834	0.926	
Asbestos Roof (1,0)	0.450	0.481	-0.031	834	0.402	
Storey (number)	1.256	1.306	-0.050	831	0.151	
House ownership (1,0)	0.764	0.785	-0.022	834	0.482	
Female (1,0)	0.209	0.183	0.026	834	0.380	
Age (years)	49.651	50.943	-1.292	834	0.142	
Married (1,0)	0.806	0.808	-0.002	834	0.953	
Education (years)	10.318	9.536	0.781	832	0.001	***
Household size (in logs)	3.702	3.689	0.013	834	0.907	
Household expenditure (million IDR)	3.769	3.294	0.474	834	0.001	***
Public health insurance (1,0)	0.826	0.836	-0.010	834	0.719	
Flood expenditure in 2017 (million IDR)	0.272	0.317	-0.045	834	0.556	

Note: This table presents the summary statistics (standard errors in brackets) and the p-values for differences in means between the randomly assigned flood-free or flood-prone groups.
Source: The Impact of Jakarta Floods Survey (2018).

The location of floodgates may be confounded by the level of flood risk, as stations might be placed in areas with higher exposure. This could potentially lead to differences in household characteristics and investment decisions between those living closer to floodgate stations and those farther away, thereby biasing insurance demand estimates. To address this concern, I conduct balance tests comparing observable household characteristics across two distinct distance bands: households residing within 5 km ("close") and beyond 19 km ("far") from their reference floodgate station. Table 2.6 shows no significant differences between the two groups in terms of housing conditions and investments. This evidence supports the

[24] The flood-free group includes households located in "almost never flooded" areas, whereas the flood-prone group consists of those situated in "occasionally flooded" and "always flooded" areas.

plausibility of the exogeneity assumption and helps address concerns that households self-select into locations based on unobserved vulnerability or investment preferences.

Table 2.6 Balance tests - distance to floodgate (close<=5 km vs. far>=19 km)

Sample: Dependent variable	**Distance: Far from and close to floodgate (km)**	
Electricity (1,0)	-0.00003	(0.0007)
Asbestos roof (1,0)	0.0005	(0.0055)
Storey (number)	0.0008	(0.0058)
House ownership (1,0)	0.0019	(0.0046)
Female (1,0)	-0.0045	(0.0046)
Age (years)	0.0847	(0.1266)
Married (1,0)	0.0037	(0.0042)
Education (years)	0.0423	(0.0362)
Household size (number)	-0.0025	(0.0163)
Household expenditure (million IDR)	-0.0279	(0.0236)
Public health insurance (1,0)	-0.0051	(0.0046)
Flood expenditure in 2017 (million IDR)	0.0046	(0.0047)
Distance to near river (in logs)	Yes	
Region FE	Yes	
Number of observations (N)	230	

Notes: OLS regressions. Each coefficient represents the result of a regression of the dependent variable on distance to the reference floodgate station. While the distance to the floodgate spans from 0.8 km to 24.4 km, the analysis focuses on a restricted sample. Specifically, the "Far from and close to floodgate station" sample includes only households located within 5 km ("close") or beyond 19 km ("far") of their reference floodgate station. Each regression includes distance from the surveyed household to nearest river and region fixed effects. Asterisk (*), double asterisk (**), and triple asterisk (***) denote variables significant at 10 percent, 5 percent, and 1 percent, respectively.
Source: The Impact of Jakarta Floods Survey (2018).

A potential concern of self-selection into flood-free and flood-prone areas based on risk preferences can cause bias in estimating the effect of distance to the reference floodgate station on insurance demand. To assess this, I test whether households' risk preferences predict their choice of residential location with respect to proximity to the floodgate and exposure to flood risk. Specifically, I regress distance to the floodgate station, flood intensity, and flood shocks on both risk preferences (from the hypothetical lottery) and extreme risk aversion. Table 2.7 shows that neither of them is statistically significant across any of the location variables. The evidence suggests that risk preferences do not predict respondents' location choices (i.e., residential sorting) with respect to distance to the floodgate and flood exposure, addressing concerns about self-selection bias in the analysis.

Table 2.7 Risk preferences and flood risks

	Dependent variable:		
Model:	(1) Distance to floodgate (km)	(2) Flood intensity (category)	(3) Flood shocks (number)
	(OLS)	(Ordered)	(OLS)
Risk preference (hypo. lottery)	0.049	-0.044	-0.002
	(0.094)	(0.041)	(0.060)
Extreme risk-aversion (1,0)	0.307	-0.105	-0.117
	(0.347)	(0.150)	(0.221)
Distance to nearest river (in logs)	Yes	Yes	Yes
Region FE	Yes	Yes	Yes
Number of observations	836	836	836

Notes: Columns 1 and 3 represent OLS regressions and column 2 an Ordered regression of measures of risk aversion on distance to the reference floodgate station, flood intensity and flood shocks. Each result displays the coefficients related to each explanatory variable. Each regression includes distance from the surveyed household to nearest river and region fixed effects. Asterisk (*), double asterisk (**), and triple asterisk (***) denote variables significant at 10 percent, 5 percent, and 1 percent, respectively. Source: The Impact of Jakarta Floods Survey (2018).

2.4.3 Empirical model

The cross-sectional equation of interest is:

$$P(y_{ij} = 1) = F(\beta_0 + \boldsymbol{\beta_1 D_i} + \boldsymbol{\beta_2 P_i} + \boldsymbol{\beta_3 E_i} + \delta G_{ij} + \gamma X_i + \lambda F_i + \phi_r + \theta_j + \varepsilon_{i,j}) \quad (9)$$

where y_{ij} is the household participation in the index-based flood insurance product (a binary variable = 1 if the ith household in the urban village j is willing to purchase the product and =0 otherwise). $\boldsymbol{D_i}$ indicates the distance of a house to the reference floodgate station expressed in km; $\boldsymbol{P_i}$ captures the premium rates of index-based flood insurance offered to the ith household; $\boldsymbol{E_i}$ represents the extremely risk averse household (a binary variable = 1 if the ith household is the most risk averse and =0 otherwise); G_j is the vector of geographical characteristics in a urban village; X_i represents the household characteristics; F_i is a vector of flood experience of the ith household; ϕ_r a fixed effect for the five regions; θ_j is the flood intensity fixed effect at urban village level; and $\varepsilon_{i,j}$ is the random error. To account for the dichotomous nature of the dependent variable, I will estimate equation (9) using probit regressions with a total sample size of 1,672 observations. The analysis controls for the distance of a house to the nearest river, as well as region-specific and flood intensity fixed effects.

I also investigate the interaction of distance to the floodgate station with extreme risk aversion, and with premium in separate regressions to identify the channels through which basis risk impacts demand for index-based flood insurance.

2.5 Results

2.5.1 Main results

In this section, I present the results on the impact of basis risk, premium and extreme risk aversion on the demand for index-based flood insurance. The probit estimations of index-based flood insurance demand for "all sample" includes an initial and follow-up bidding questions asked to surveyed households. Additionally, I present separately these estimations for responses to the "first question" and "second question" in order to identify whether or not I face potential problems of sequential bidding experiments influenced by their starting point (Cameron and James, 1987). Finally, I bootstrap all probit regressions (1,000 replications) to approximate standard errors, confidence intervals, and p-values for test statistics, based on the sample data[25].

Based on equation (9), in columns (1) through (3) of Panel A Table 2.8, results suggest that demand for index-based flood insurance is negative and statistically significant at 1 percent for distance to the floodgate station and premium, and negative and statistically significant at 10 percent for extremely risk-averse households, as predicted. When using probit regressions for double-bounded dichotomous choice ("all sample"), and separate first and second questions, I find no significant differences in terms of coefficients and significance level. However, through "first" and "second" question specifications, extremely risk averse yields non-significant estimates. Additionally, a 10 percent decrease in price seems to lead to a 1.3 percentage point increase in uptake which, given the effective demand of 9.9 percent, corresponds to a 12.8 percent increase in demand for this product[26].

[25] The sample size, while sufficient for the analysis, may not be large enough for traditional asymptotic approximations to be fully accurate. Bootstrapping helps to mitigate this issue by generating an empirical distribution of the estimator from the data.

[26] This is calculated with the product of the average marginal effect for the premium (in logs) variable in column (1) of Table 3, that is -0.133, and the natural logarithm transformation of a 10 percent increase in the same variable, which is *ln*(1.1).

Table 2.8 Index-based flood insurance demand among sampled households

	Dependent variable: Insurance demand					
Model: Probit	(1) All sample	(2) 1st question	(3) 2nd question	(4) All sample	(5) All sample	(6) Far from and close to floodgate
PANEL A						
Distance to floodgate (km)	-0.013 ***	-0.015 ***	-0.010 **	-0.015 ***	-0.032 ***	
	(0.003)	(0.004)	(0.004)	(0.003)	(0.007)	
Premium (in logs)	-0.133 ***	-0.152 ***	-0.106 ***	-0.133	-0.179 ***	-0.093 ***
	(0.008)	(0.009)	(0.015)	(0.008)	(0.017)	(0.015)
Extreme risk-aversion (1,0)	-0.038 **	-0.029	-0.042	-0.094	-0.041 **	
	(0.019)	(0.028)	(0.030)	(0.039)	(0.020)	
Distance to floodgate (km) x Extreme risk-aversion (1,0)				0.006		
				(0.003)		
Distance to floodgate (km) x Premium (in logs)					0.005 ***	
					(0.002)	
Floodgate station is close (≤ 5km)						0.385 ***
						(0.113)
Floodgate station is close (≤ 5km) x Premium (in logs)						-0.076 ***
						(0.027)
PANEL B						
Bias-adjusted estimates, with Oster δ=1 and $Rmax$=1.3$\tilde{R}$						
- Distance to floodgate station (km)	-0.021	0.024	-0.016	-	-	-
- Premium (in logs)	-0.148	-0.184	-0.113	-	-	-
- Extreme risk aversion (1,0)	-0.023	-0.015	-0.029	-	-	-
Distance to nearest river (in logs)	Yes	Yes	Yes	Yes	Yes	Yes
Region FE	Yes	Yes	Yes	Yes	Yes	Yes
Flood intensity FE	Yes	Yes	Yes	Yes	Yes	Yes
Covariates	Yes	Yes	Yes	Yes	Yes	Yes
Number of observations	1,662	831	828	1,662	1,662	910

Notes: Panel A presents probit regression estimations of household-level insurance uptake on distance to the reference floodgate station, premium and extreme risk-aversion. Average marginal effects are reported. All sample corresponds to double-bounded dichotomous choice questions, except for columns (1) and (2). The "Far from and close to floodgate station'' sample only includes households at a perpendicular distance from their reference floodgate station below 5 km (''close'') or over 12 km (''far''). In columns (4), (5) and (6) probit specifications include the interaction term. Panel B shows Oster test results. All specifications include distance to nearest river, covariates, and region and flood intensity fixed effects. Standard errors are bootstrapped (1,000 replications) and reported in parentheses. Asterisk (*), double asterisk (**), and triple asterisk (***) denote variables significant at 10 percent, 5 percent, and 1 percent, respectively. Source: The Impact of Jakarta Floods Survey (2018).

I now turn to the relationship between basis risk and extreme risk aversion. Specifically, I expect the extremely risk-averse households to be more likely to purchase the insurance product when they are closer to the reference floodgate station than far away. Although the coefficient of the interaction term in column (4) is not statistically significant, it has the expected sign (see Figure 2.A.3 in the Appendix).

In column (5), I only consider the effects of basis risk and premium. I assume there is different price elasticity of insurance that households face as they move closer (less basis risk) or further (more basis risk) from the reference floodgate station. I test this assumption by interacting premium and distance to the floodgate station. I find a strong significance estimate resulting in price elasticity increases the closer the household is to the reference floodgate station.

As an additional exercise, in column (6) of Table 2.8, the sample is restricted to only those households located less than 5 kilometres (km) and more than 12 km from their reference floodgate station[27]. I then use an indicator variable that takes the value of one if a household belongs to the first group (less than 5 km), and otherwise zero. Then, doubling the distance of a house to the reference floodgate station reduces insurance demand by 0.9 percentage points.

---

[27] The idea to restrict the sample comes from the study by Hill et al. (2016) that considers the average distance to the reference station of the treated (5km) and control (10km) villages in India. In the same way, I find the same difference in distance to the floodgate station variable where, on average, households that always experience floods are 12km from the reference floodgate station while households that almost never or occasionally experience floods are 8km from the reference station; therefore, I sample for 5km ("close") and 12km ("far").

Table 2.9 Index-based flood insurance demand among sampled households using Linea Probability Model (LPM + 1)

	Dependent variable: Insurance demand											
Model: LPM +1	(1) All sample		(2) 1st question		(3) 2nd question		(4) All sample		(5) All sample		(6) Far from and close to floodgate	
Distance to floodgate (km)	-0.012	***	-0.014	***	-0.010	***	-0.014	***	-0.037	***		
	(0.002)		(0.003)		(0.003)		(0.003)		(0.008)			
Premium (in logs)	-0.146	***	-0.183	***	-0.112	***	-0.146		-0.201	***	-0.100	***
	(0.009)		(0.011)		(0.015)		(0.009)		(0.018)		(0.015)	
Extreme risk-aversion (1,0)	-0.030	*	-0.025		-0.033		-0.081	**	-0.033	*		
	(0.018)		(0.026)		(0.026)		(0.037)		(0.018)			
Distance to floodgate (km) x Extreme risk-aversion (1,0)							0.005					
							(0.003)					
Distance to floodgate (km) x Premium (in logs)									0.006	***		
									(0.002)			
Floodgate station is close (≤ 5km)											0.459	***
											(0.120)	
Floodgate station is close (≤ 5km) x Premium (in logs)											-0.085	***
											(0.024)	
Distance to nearest river (in logs)	Yes		Yes		Yes		Yes		Yes		Yes	
Region FE	Yes		Yes		Yes		Yes		Yes		Yes	
Flood intensity FE	Yes		Yes		Yes		Yes		Yes		Yes	
Covariates	Yes		Yes		Yes		Yes		Yes		Yes	
Number of observations	1,662		831		831		1,662		1,662		910	

Notes: This table presents Linear Probability Model regressions of household-level insurance uptake on distance to the reference floodgate station, premium and extreme risk-aversion. All samples correspond to double-bounded dichotomous choice questions, except for columns (1) and (2). The "Far from and close to floodgate station'' sample only includes households at a perpendicular distance from their reference floodgate station below 5 km (''close'') or over 12 km (''far''). In columns (4), (5) and (6), LPM +1 specifications include the interaction term. All specifications include distance to nearest river, covariates, as well as region and flood intensity fixed effects. Standard errors are bootstrapped (1,000 replications) and reported in parentheses. Asterisk (*), double asterisk (**), and triple asterisk (***) denote variables significant at 10 percent, 5 percent, and 1 percent, respectively.
Source: The Impact of Jakarta Floods Survey (2018).

To verify the robustness of the main results, I employ the method developed by Oster (2019) to understand whether variation in unobservables could drive the results – especially in the case of a proxy for basis risk (see Panel B of Table 2.8).

I set the coefficient of proportionality equal to one (δ=1) which suggests that "the observables are at least as important as the unobservables" (Oster, 2019: 195-196) in determining the outcome variable. In addition, I set the value of R^2_{max}, the R^2 from a hypothetical regression of the outcome on basis risk and both observed and unobserved controls, to be equal to $1.3\tilde{R}^2$, where $\tilde{R}^2$is the R^2 from the corresponding regression from Panel A of Table 2.8. The magnitude of Oster's statistics makes it very unlikely that the results of distance of a house to the floodgate station can be explained by variation in unobservables.

Moreover, Table 2.9 reports the Linear Probability Model (+1 LPM) results which are relatively the same in magnitude and signs as the main estimates[28].

[28] This is also true for the risk preferences with real monetary payoffs instead of hypothetical rewards (see Table 2.A.2 in the Appendix).

Table 2.10 Index-based flood insurance demand among sampled households excluding covariates

	Dependent variable: Insurance demand					
Model: Probit	(1) All sample	(2) 1st question	(3) 2nd question	(4) All sample	(5) All sample	(6) Far from and close to floodgate
Distance to floodgate (km)	-0.006***	-0.006**	-0.005*	-0.007***	-0.025***	
	(0.002)	(0.003)	(0.003)	(0.002)	(0.007)	
Premium (in logs)	-0.131***	-0.152***	-0.105***	-0.131***	-0.179***	-0.087***
	(0.008)	(0.009)	(0.014)	(0.008)	(0.018)	(0.013)
Extreme risk-aversion (1,0)	-0.054***	-0.048*	-0.057**	-0.113***	-0.056***	
	(0.020)	(0.027)	(0.028)	(0.041)	(0.019)	
Distance to floodgate (km) x Extreme risk-aversion (1,0)				0.006		
				(0.004)		
Distance to floodgate (km) x Premium (in logs)					0.005***	
					(0.002)	
Floodgate station is close (≤ 5km)						0.317***
						(0.105)
Floodgate station is close (≤ 5km) x Premium (in logs)						-0.079***
						(0.026)
Distance to nearest river (in logs)	Yes	Yes	Yes	Yes	Yes	Yes
Region FE	Yes	Yes	Yes	Yes	Yes	Yes
Flood intensity FE	Yes	Yes	Yes	Yes	Yes	Yes
Covariates	Yes	Yes	Yes	Yes	Yes	Yes
Number of observations	1,672	836	836	1,672	1,672	914

Notes: This table presents probit regression estimations of household-level insurance uptake on distance to the reference floodgate station, premium and extreme risk-aversion, excluding covariates. Average marginal effects are reported. All sample corresponds to double-bounded dichotomous choice questions, except for columns (1) and (2). In columns (4), (5), and (6) probit specifications include the interaction term. All specifications include distance to nearest river, as well as region and flood intensity fixed effects. Standard errors are bootstrapped (1,000 replications) and reported in parentheses. Asterisk (*), double asterisk (**), and triple asterisk (***) denote variables significant at 10 percent, 5 percent, and 1 percent, respectively.
Source: The Impact of Jakarta Floods Survey (2018).

Furthermore, I first exclude the set of covariates from the estimations. The results in Table 2.10 show that all point estimates remain significant and relatively similar. This can be interpreted as evidence of structural validity of the probit specification.

Second, I use the available time-series data of floods from BPBD to perform a placebo test for the identification assumption with respect to basis risk. Specifically, I focus on urban villages that experienced flood water levels above 2 metres registered between 2013 and 2017[29], as these areas are typically prone to flooding during the rainy season. Then, I designate the centre of each urban village as the hypothetical floodgate station[30]. Using this approach, I perform the placebo test by measuring the distance from each surveyed household to the urban village centre (i.e., the hypothetical floodgate location)[31] to examine the validity of the identification assumption.

As expected, placebo estimates are non-significant (Table 2.11), and close to zero. This validates that there are no other distances from a house to another reference flood point that have negative effects on index-based flood insurance demand.

Finally, given the sufficient price variation along the hypothetical downward-sloping demand curve, I redo the regression in equation (9) for different levels of premium. Overall, it seems that households who face premiums below IDR 50,000 (USD 3.45) are more likely to purchase the insurance product; conversely, they are unwilling to buy when the premium is above this value (see Table 2.A.3 in the Appendix). Additionally, the mean monthly willingness-to-pay (MWTP) is estimated at IDR 8,299.65 (USD 0.57), representing the average amount households in Jakarta are willing to pay each month for index-based flood insurance (see Table 2.A.4). This amount represents 4.97% of the total insurance premium.

---

[29] A box plot is used to visually identify outliers within the flood water level data set, defined as points that are located outside the whiskers of the box plot. These points lie on the range of 200 cm – 400 cm and provide the information of urban villages that experienced these flood heights (see Figure 2.A.4 in the Appendix).

[30] I use urban village centres since they are often points of reference for local governance, infrastructure, and community organisation, and are typically areas of high relevance to flood response strategies.

[31] Using the OpenStreetMap website to obtain x-y coordinates of the centroid of these urban villages, I calculated the perpendicular distance of respondents' houses to the hypothetical floodgate station.

Table 2.11 Distance to hypothetical floodgate stations located in urban village centres with flood levels above 2m

Model: Probit	Dependent variable: Insurance demand (1) All sample		(2) 1st question		(3) 2nd question		(4) All sample		(5) All sample		(6) Far from and close to floodgate	
Distance to hypothetical floodgate (km)	-0.0003		-0.017	***	-0.00003		-0.0002		0.0089			
	(0.008)		(0.004)		(0.007)		(0.007)		(0.012)			
Premium (in logs)	-0.131	***	-0.152	***	-0.103	***	-0.131	***	-0.115	***	-0.142	***
	(0.011)		(0.009)		(0.017)		(0.009)		(0.014)		(0.029)	
Extreme risk-aversion (1,0)	-0.046	**	-0.028		-0.047		0.012		-0.046	**		
	(0.022)		(0.028)		(0.031)		(0.031)		(0.021)			
Distance to hypothetical floodgate (km) x Extreme risk-aversion (1,0)							-0.010	**				
							(0.004)					
Distance to hypothetical floodgate (km) x Premium (in logs)									-0.003			
									(0.002)			
Floodgate station is close (≤ 5km)											0.039	
											(0.145)	
Floodgate station is close (≤ 5km) x Premium (in logs)											0.018	
											(0.033)	
Distance to nearest river (in logs)	Yes		Yes		Yes		Yes		Yes		Yes	
Region FE	Yes		Yes		Yes		Yes		Yes		Yes	
Flood intensity FE	Yes		Yes		Yes		Yes		Yes		Yes	
Covariates	Yes		Yes		Yes		Yes		Yes		Yes	
Number of observations	1,662		831		828		1,662		1,662		1,208	

Notes: This table presents probit regressions of index-based flood insurance uptake on distance to the hypothetical floodgate station, premium rates and extreme risk aversion. Distance to the hypothetical floodgate station is represented by the distance of a house to the closest area with flood level above 2 metres which is the placebo. Average marginal effects are reported. All sample corresponds to double-bounded dichotomous choice questions, except for columns (1) and (2). In columns (4), (5) and (6) probit specifications include the interaction term. All probit regressions include distance to nearest river, covariates, as well as region and flood intensity fixed effects. Standard errors are bootstrapped (1,000 replications) and reported in parentheses. Asterisk (*), double asterisk (**), and triple asterisk (***) denote variables significant at 10 percent, 5 percent, and 1 percent.
Source: The Impact of Jakarta Floods Survey (2018).

2.5.2 Extended analysis

In this section, I extend the investigation by looking at the heterogeneity of household responses to basis risk, premiums and risk aversion. First, I test whether the index-based flood insurance product demand responds to price changes through different degrees of basis risk. The intuition behind this relationship is: i) when insurance has a low basis risk (close to zero) and is at an actuarially-fair price ($m = 1$), the price elasticity demand is relatively high. This means households are more likely to purchase insurance, particularly if they receive price discounts; and, ii) when insurance has a high basis risk for any amount at or below the actuarially-fair price ($m \leq 1$), the demand price elasticity goes closer towards zero.

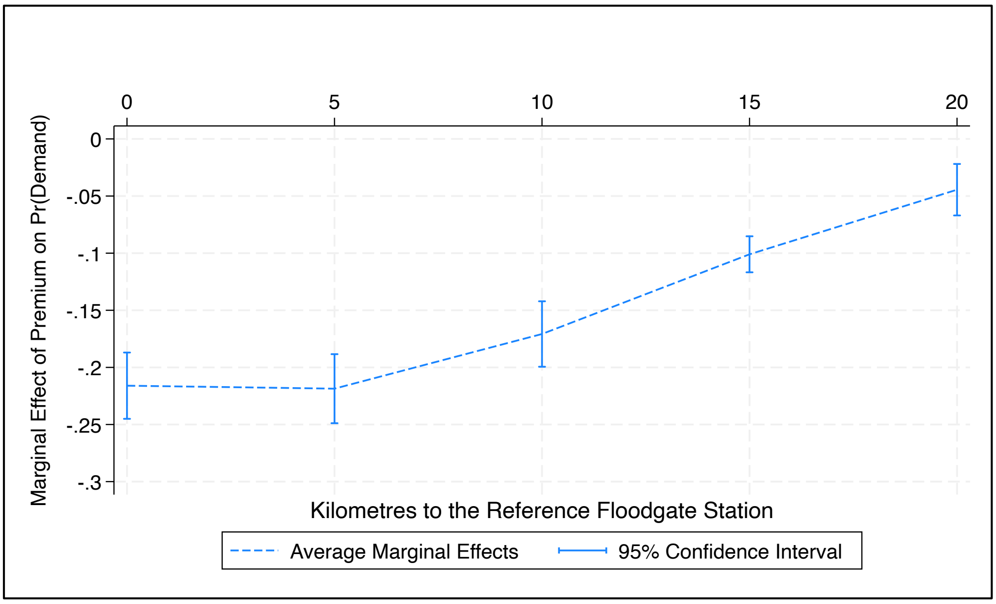


Figure 2.5 Price sensitivity of demand and distance to reference floodgate station

Notes: This figure shows the relationship between estimated price sensitivity of insurance demand and distance to the reference floodgate station. The dashed line plots average marginal effects of price on insurance uptake, conditional on being located a certain distance from the reference floodgate station. The solid lines plot 95 percent confidence bands around the average marginal effect point estimates. The figure stems from the estimation of a probit regression model of household-level insurance uptake, on variations of prices, distance of the household location to the reference floodgate station, and the interaction between the two, including covariates. This specification includes distance to nearest river, as well as region and flood intensity fixed effects.
Source: The Impact of Jakarta Floods Survey (2018).

Figure 2.5 shows the average marginal effect of the logarithm of price on the probability of insurance uptake for households located at different distances from the reference floodgate station. Households located less than 5 km from a reference station have a sensitivity to price 4.9 times higher (marginal effect of –0.22) than those located more than 12 km from a floodgate station (marginal effect of –0.04). Based on the above findings, I can conclude that sensitivity to price discounts increase uptake when floodgate stations closer to households are targeted to reduce basis risk. This is an important finding given the potential subsidies (via discounts) on flood insurance contracts.

Second, I turn to testing whether demand for index insurance varies across difference levels of risk aversion. The Clarke's (2016) model discussed in the theoretical framework section indicates that, in the presence of basis risk ($r > 0$), when premiums are below actuarially-fair, a downward demand with respect to risk aversion is expected. I explicitly test for this prediction and find results consistent with the above: for an index insurance product with a multiple at or below 1 ($m \leq 1$), the intensity of demand is negatively related to risk aversion.

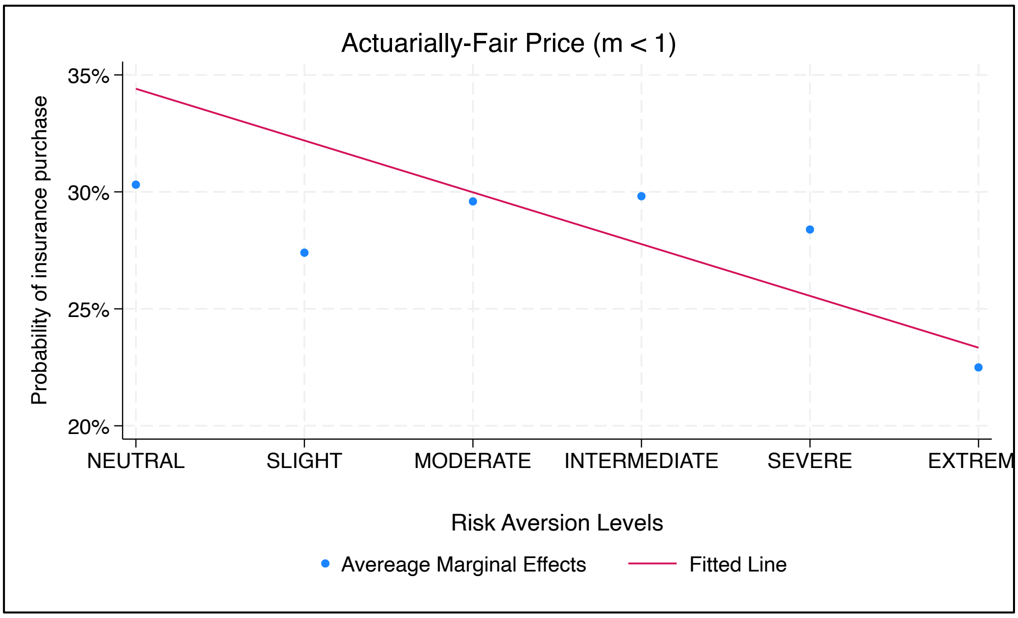


Figure 2.6 Probability of insurance purchase across levels of risk aversion

Notes: This figure shows the estimated probability of purchasing insurance against different levels of risk aversion. The point estimates are average marginal effects of risk aversion levels on insurance uptake. The solid line plots the fitted line that crosses all the point estimates and shows the downward-slope relationship between risk aversion and demand for an insurance product with a multiple at or below 1. The figure stems from the estimation of a probit regression model of household-level insurance uptake and risk aversion levels, including covariates. This specification includes distance to nearest river (in logs), as well as region and flood intensity fixed effects.
Source: The Impact of Jakarta Floods Survey (2018).

Using historical weather data from BPBD, I estimate the actuarially-fair price of the hypothetical index insurance that was offered (set at IDR 167,000). Households randomly assigned a monthly premium of IDR 100,000 or less, offered a 50% discount following a negative response to the initial question, and located within 5 km of the reference floodgate[32],

[32] This resulted in a group of surveyed households representing 8.49% of the total sample.

faced actuarially favourable hypothetical insurance (with a multiple equal to and less than 1)[33]. The predicted downward-sloping demand curve for urban households under a favourable index insurance product is shown in Figure 2.6 The probability of purchasing index insurance for the least risk-averse households is 7.8 percentage points higher than for those who are extremely risk-averse.

In addition, I look at how the price sensitivity of demand varies across different levels of risk aversion. For this reason, I use the following probit specification:

$$P\left(y_{ij} = 1\right) = F(\alpha_0 + \alpha_{\mathbf{1}}\boldsymbol{P_i} + \alpha_{\mathbf{2}}\boldsymbol{R_i} + \alpha_{\mathbf{3}}\boldsymbol{P_i} * \boldsymbol{R_i} + \delta G_{ij} + \gamma X_i + \lambda F_i + \phi_r + \theta_j + \varepsilon_{i,j}) \quad (10)$$

where I exclude distance to the reference floodgate station $\boldsymbol{D_i}$ and $\boldsymbol{E_i}$ from equation (9) and introduce $\boldsymbol{R_i}$ which represents the level of risk aversion of ith household, along with the vector of geographical and household characteristics, fixed effects, and random error.

I expect the price elasticity of demand for insurance to be higher among households who are least risk averse than those who are extreme risk averse. I plot the average marginal effect of logarithm of price on the probability of uptake for households with different levels of risk aversion (Figure 2.7). I observe that least risk averse households have a sensitivity price 1.1 times higher than most risk averse households. Overall, this result appears consistent with the theoretical prediction.

[33] Closer to the reference floodgate, index insurance still presents basis risk, but to a lesser extent than in more distant areas, resulting in a more accurate flood index. As a result, m is more likely to be ≤ 1, indicating better alignment between the premium and the expected losses. I choose a 5 km distance from the reference floodgate because the results show that demand is significantly more price-sensitive nearby than in more distant areas.

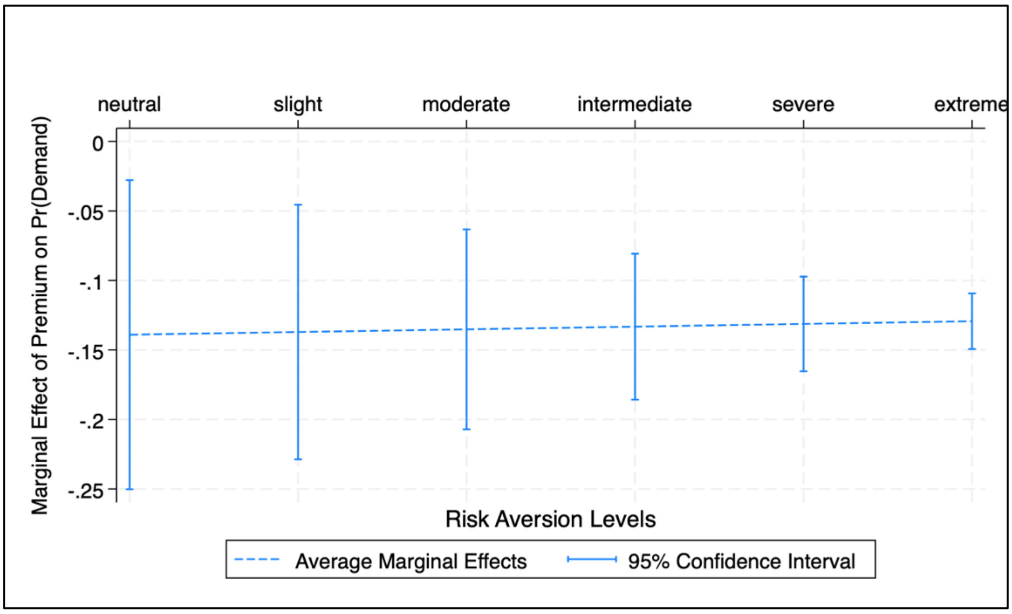


Figure 2.7 Price sensitivity across levels of risk aversion

Notes: This figure shows the probability of insurance purchase against different levels of risk aversion. The dashed line plots average marginal effects of price on insurance uptake. The solid lines plot 95 percent confidence bands around the average marginal effect point estimates. The figure stems from the estimation of a probit regression model of household-level insurance uptake, on variations of prices, extreme risk aversion, and the interaction between the two, including covariates. This specification includes distance to nearest river, as well as region and flood intensity fixed effects.
Source: The Impact of Jakarta Floods Survey (2018).

Third, this study examines the price sensitivity of demand in each region[34]. I test the assumption that price sensitivity increases at lower levels of basis risk – i.e., it is higher closer to the floodgate station and lower far away. I observe the point estimates for interaction terms with positive sign and significance levels at 10% and 5% on West and South regions, respectively, as shown in Table 2.12. These findings tell us that basis risk is problematic in these two regions. This may be explained by the small number of stations compared to other regions, and the average distance, 7 km and 13 km, between the household and the reference floodgate station. Overall, houses located in the range between 0 km and 5 km seem to estimate the degree of geographical variation in floods and loss occurring at their location as being similar to what may be recorded at the floodgate station.

[34] I exclude Central Jakarta region due to the small number of respondents in the sample (17 households) which would not allow us to regress the probit specification. Also, this region is the least vulnerable from flooding impacts (with an average flood height of 40 cm within the 2013-2017 period) due to better flood mitigation infrastructure situated around the city centre.

Table 2.12 Price sensitivity demand and distance to floodgate station in each region

	Dependent variable: Insurance demand							
Model: Probit	(1) West Jakarta		(2) South Jakarta		(3) East Jakarta		(4) North Jakarta	
Distance to floodgate (km)	-0.1080	**	-0.0092		-0.0099		-0.0446	
	(0.053)		(0.011)		(0.019)		(6.509)	
Premium (in logs)	-0.335	***	-0.209	***	-0.129	***	-0.222	
	(0.081)		(0.028)		(0.046)		(7.107)	
Distance to floodgate (km) x Premium (in logs)	0.023	*	0.005	**	0.003		0.009	
	(0.012)		(0.002)		(0.004)		(1.097)	
Distance to nearest river (in logs)	Yes		Yes		Yes		Yes	
Flood intensity FE	Yes		Yes		Yes		Yes	
Covariates	Yes		Yes		Yes		Yes	
Number of observations	396		474		560		192	

Notes: This table presents probit regression estimations of household-level insurance uptake on distance to the reference floodgate station, premium rates and the interaction term (which is the price sensitivity demand). Average marginal effects are reported. All sample corresponds to double-bounded dichotomous choice questions. All columns include distance to nearest river, covariates, as well as flood intensity fixed effects. Standard errors are bootstrapped (1,000 replications) and reported in parentheses. Asterisk (*), double asterisk (**), and triple asterisk (***) denote variables significant at 10 percent, 5 percent, and 1 percent, respectively.
Source: The Impact of Jakarta Floods Survey (2018).

Fourth, I observe the results by subgroups of income, education level, and flood areas. Columns (1) to (6) in Table 2.13 show that, in general, results are likely not statistically different among these subgroups. It is clear that premiums are strongly significant in reducing index-based flood insurance demand across subgroups; while distance to floodgate station effects on demand have interesting results. In columns (1) through (4) of Table 2.13, households with high income and education seem to be slightly sensitive to basis risk compared to those lower in the two categories. This could be due to richer households having a lack of trust for the government to provide index-based flood insurance in Jakarta. Finally, households located in flood zone areas are more sensitive to basis risk and premium. This could be due to the poor usually settling in areas exposed to floods and low amenities, and more sensitive to price changes.

Table 2.13 Index-based flood insurance demand by different subgroups

Model: Probit	Dependent variable: Insurance demand											
	(1) Low income		(2) High income		(3) Never-attended university		(4) Attended university		(5) Within flood zone		(6) Outside flood zone	
Distance to floodgate station (km)	-0.0096	**	-0.0151	***	-0.0123	**	-0.0121	***	-0.0224	***	-0.0082	**
	(0.005)		(0.005)		(0.005)		(0.004)		(0.005)		(0.004)	
Premium (in logs)	-0.116	***	-0.151	***	-0.139	***	-0.125	***	-0.143	***	-0.113	***
	(0.012)		(0.011)		(0.012)		(0.012)		(0.010)		(0.015)	
Extremely risk-averse (1,0)	-0.072	**	-0.010		-0.047		-0.019		-0.039		-0.049	
	(0.029)		(0.030)		(0.029)		(0.029)		(0.024)		(0.038)	
Distance to nearest river (in logs)	Yes		Yes		Yes		Yes		Yes		Yes	
Region FE	Yes		Yes		Yes		Yes		Yes		Yes	
Flood intensity FE	Yes		Yes		Yes		Yes		No		No	
Covariates	Yes		Yes		Yes		Yes		Yes		Yes	
Number of observations	828		830		820		842		1,146		514	

Notes: This table presents probit regression of household-level insurance uptake on distance to the reference floodgate station, premium rates and extreme risk aversion. I sample the data set by subgroups relative to the household's income, years of education and whether the household is located within the flood zone. Average marginal effects are reported. All probit regressions include distance to nearest river, covariates, as well as region and flood intensity fixed effects. Standard errors are bootstrapped (1,000 replications) and reported in parentheses. Asterisk (*), double asterisk (**), and triple asterisk (***) denote variables significant at 10 percent, 5 percent, and 1 percent.
Source: The Impact of Jakarta Floods Survey (2018).

2.6 Conclusions

This study presents causal evidence on three factors that have effects on the uptake of index-based flood insurance in Jakarta, Indonesia: distance of a house to the reference floodgate station (a proxy of basis risk), premium rates and household risk aversion. I link an expected utility framework developed by Hill et al. (2016), complemented with elements of Clarke's (2016) work, to the empirical analysis to gain a better understanding of index-based flood insurance demand under the presence of basis risk in the urban context.

The results in this article indicate that demand for the index-based flood insurance product offered to urban households decreases as the degree of basis risk and premium increase and declines at higher levels of risk aversion due to the imperfect coverage provided by a product with basis risk. These effects are robust and significant.

In order to design an improved insurance product for the future, it is important to consider the presence of basis risk. The results find that households are more likely to purchase the insurance product when they are closer to the reference floodgate station, i.e., insurance product with lower basis risk. In addition, they verify the price elasticity of demand for insurance is higher when basis risk is low and among households who are less risk averse. Doubling the probability of having basis risk reduces insurance demand by 0.9 percentage points.

The general lessons learned from this study for urban areas in developing countries are as follows. It is important to recognise that the underlying ground station data is generally sparse in developing countries (Dell et al., 2014), and flooding is typically identified as spatially correlated shocks. Having enough floodgate stations so as to reduce the degree of basis risk and encourage future demand is key for successfully implementing index-based flood insurance. The rule of thumb concluded from the analysis is that approximately 5 km should be the maximum distance between any house and a floodgate station. If this is not the case, it is integral to invest in the development of new floodgates prior to introducing the insurance product.

Reliable and detailed statistical information from floodgate stations provides the data and hazard modelling necessary to assess the distribution of flood-related losses over time and across space. Such information enables the design of index insurance products that are more closely aligned with local damages. Investments in data infrastructure generate benefits that can be shared across multiple insurance contracts, meaning that once established, each household's contribution to the costs is relatively small. Moreover, local governments can leverage this data to offer index insurance schemes with lower basis risk, increasing uptake among households seeking protection against flood hazards. Complementary measures—such as sessions introducing potential flood-related risks households might face, discussions on coping mechanisms and communication about the design and benefits of index insurance—are equally important to strengthen confidence in the product. Implementing these measures enhances the potential for broader adoption of insurance instruments to improve disaster risk financing, particularly for community-level flood protection in Jakarta, Indonesia.

In line with the effect on demand of reduced insurance prices, I recommend the application of different levels of premium discounts depending on household location. The results show that subsidies have an immediate effect on insurance demand, especially households who are offered insurance products with lower basis risk. However, the exact design and structure of a premium discount offering is still unclear in this study due to the unknown nature of household welfare, particularly regarding the level of discount needed to incentivise insurance uptake. This issue presents an opportunity for further research.

References

Barnett, BJ, Mahul, O 2007, ‘Weather index insurance for agriculture and rural areas in lower-income countries’, *American Journal Agriculture Economics*, vol. 89, issue 5, pp. 1241–1247.

Barnett, BJ, Barrett, CB, Skees, JR 2008, ‘Poverty traps and index-based risk transfer products’, *World Development*, vol. 36, issue 10, pp. 1766-1785.

Binswanger, H 1980, ‘Attitudes toward risk: experimental measurement in rural India’, *American Journal of Agricultural Economics*, vol. 62, no. 3, pp. 395–407.

Bauer, M, Chytilová, J, Morduch, J 2012, ‘Behavioral Foundations of Microcredit: Experimental and Survey Evidence from Rural, India’, American Economic Review, vol. 102, no. 2, pp. 1118-1139.

Ceballos, F, Robles, M 2020, ‘Demand heterogeneity for index-based insurance: the case for flexible products’, *Journal of Development Economics*, vol. 146.

Chowdhury, S, Sutter, M and Zimmermann, K 2022, ‘Economic Preferences across Generations and Family Clusters: A Large-Scale Experiment in a Developing Country’, *Journal of Political Economy*, vol, 130, no. 9, pp. 2361-2410.

Chuang, Y, Schechter, L 2015, ‘Stability of experimental and survey measures of risk, time, and social preferences: a review and some new results’, *Journal of Development Economics*, vol. 117, pp. 151-70.

Clarke, DJ 2016, ‘A theory of rational demand for index insurance’, *American Economic Journal: Microeconomics*, vol. 8, no. 1, pp. 283–306.

Cobian, J, Resosudarmo, BP 2019, ‘The cost of floods in developing countries’ megacities: a hedonic price analysis of the Jakarta housing market, Indonesia’, *Environmental Economics and Policy Studies*, no. 21, pp. 555-577.

Cameron, TA, James, MD 1987, ‘Efficient estimation methods for ‘closed-ended’ contingent valuation surveys’ *The Review of Economics and Statistics*, vol. 69, no. 2, pp. 269-276.

Cole, S, Giné, X, Tobacman, J, Topalova, P, Townsend, R, Vickery, J 2013, 'Barriers to household risk management: evidence from India', *American Economic Journal: Applied Economics*, vol. 5, no. 1, pp 104–135.

Dell, M, Jones, BF, Olken, BA 2014, 'What do we learn from the weather? the new climate-economy literature', *Journal Economic Literature*, vol. 52, no. 3, pp. 740–98.

Deng, X, Barnett, BJ, Vedenov, DV, West, JW 2007, 'Hedging dairy production losses using weather-based index insurance', *American Journal of Agricultural Economics*, no. 36, pp. 271–280.

Eckel, CC, Grossman, PJ 2002. 'Sex differences and statistical stereotyping in attitudes toward financial risk', *Evolution and Human Behaviour*, vol. 23 no. 4, pp. 281–295.

Firman, T, Surbakti, IM, Idroes, IC, Simarmata, HA 2011, 'Potential climate-change related vulnerabilities in Jakarta: challenges and current status', *Habitat International*, no. 35, pp. 372–378.

Gaurav, S, Chaudhary, V 2020, 'Do farmers care about basis risk? Evidence from a field experiment in India', *Climate Risk Management*, no. 27.

Giné, X, Townsend, R, Vickery, J 2008, 'Patterns of rainfall insurance participation in rural India' *The World Bank Economic Review*, no. 22, pp.539-566.

Hanemann, M 1991, 'Willingness to pay and willingness to accept: how much can they differ?', *American Economic Review,* vol. 81, issue 3, pp. 635-47.

Hill, RV, Hoddinott, J, Kumar, N 2013, 'Adoption of weather-index insurance: learning from willingness to pay among a panel of households in rural Ethiopia', *American Journal of Agricultural Economics*, vol. 44, no. 4-5, pp. 385–398.

Hill, RV, Robles, M, Ceballos, F 2016, 'Demand for a simple weather insurance product in India: theory and evidence', *American Journal of Agricultural Economics*, vol. 98, pp. 1250-1270.

IFAD, International Fund for Agricultural Development 2010, 'The potential for scale and sustainability in weather index insurance for agriculture and rural livelihoods', International Fund for Agricultural Development and World Food Program.

IPCC, Intergovernmental Panel on Climate Change 2012, 'Managing the risks of extreme events and disasters to advance climate change adaptation', Cambridge University Press.

-- 2018, 'Summary for policy makers: global warming of 1.5°c', *Sustainable Development, and Efforts to Eradicate Poverty*, <https://www.ipcc.ch/sr15/>.

Jha, AK, Bloch, R, Lamond, J 2012, *Cities and flooding a guide to integrated urban flood risk management for the 21st century*, World Bank, Washington.

Jensen, ND, Barrett, CB, Mude, AG 2016, 'Index insurance quality and basis risk: evidence from northern Kenya', *American Journal of Agricultural Economics*, vol. 98, no. 5, pp. 1450-1469.

Jensen, ND, Mude, AG, Barrett, CB 2017, 'How basis risk and spatiotemporal adverse selection influence demand for index insurance: evidence from northern Kenya', *Social Science Research Network*, <https://ssrn.com/abstract=2475187>.

Karadja, M, Prawitz, E 2017, 'Exit, voice, and political change: evidence from Swedish mass migration to the United States', *Journal of Political Economy*, vol. 127, no. 4, pp. 1864-1925.

Kousky, C 2019, 'The role of natural disaster insurance in recovery and risk reduction', *Annual Review of Resource Economics*, no. 11, pp. 3.1-3.20.

McIntosh, C, Sarris, A, Papadopoulos, F 2013, 'Productivity, credit, risk, and the demand for weather index insurance in smallholder agriculture in Ethiopia', *American Journal of Agricultural Economics*, no. 44, pp. 339–417

Mobarak, AM, Rosenzweig, M 2013, 'Informal risk sharing, index insurance, and risk taking in developing countries', *American Economic Review*, vol. 103, no. 3, pp. 375-380.

Octavianti, T, Charles, K 2018, 'The evolution of Jakarta's flood policy over the past 400 years: The lock-in of infrastructural solutions' *Environment and Planning C: Politics and Space*, pp. 1-23.

Oster, Emily. 2019. 'Unobservable selection and coefficient stability: theory and evidence', *Journal of Business & Economic Statistics*, vol. 37, no. 2, pp. 187-204.

Schildberg-Hörisch, H 2018, ‘Are risk preferences stable?’, *Journal of Economic Perspectives*, vol. 32, no. 2, pp. 135-154.

Sedlar, F 2016, ‘Inundated infrastructure: Jakarta’s failing hydraulic infrastructure’, *Michigan Journal of Sustainability*, vol. 4, pp. 33-45.

Skees, JR, Barnett, BJ 2006, ‘Enhancing microfinance using index-based risk-transfer products’, *Agricultural Finance Review*, vol. 66, no. 2, pp. 235-250.

Sibiko, KW, Veettil, PC, Qaim, M 2018, ‘Small farmers’ preferences for weather index insurance: insights from Kenya’, *Agriculture and Food Security*, pp. 7-53.

Soetanto, R, Hermawan, F, Milne, A, Hatmoko, JUD, As'ad, S, He C 2020, ‘Developing sustainable arrangements for “proactive” disaster risk financing in Java, Indonesia’, *International Journal of Disaster Resilience in the Built Environment*, vol. 11, pp. 435-451.

Tadesse, MA, Alfnes, F, Erenstein, O, Holden, ST 2017, ‘Demand for a labour-based drought insurance scheme in Ethiopia: a stated choice experiment approach’, *American Journal of Agricultural Economics*, vol. 48, pp. 501–511.

Takahashi, K, Ikegami, M, Sheahan, M, Barrett, CB 2016, ‘Experimental evidence on the drivers of index-based livestock insurance demand in Southern Ethiopia’, *World Development*, no. 78, pp. 324–340.

Texier, P 2008, ‘Floods in Jakarta: when the extreme reveals daily structural constraints and mismanagement’, *Disaster Prevention and Management*, vol. 17, no. 3, pp. 358-372.

United Nations Office for Disaster Risk Reduction 2019, Global Assessment Report on Disaster Risk Reduction, United Nations International Strategy for Disaster Reduction, Geneva, Switzerland, <https://www.undrr.org/publication/global-assessment-report-disaster-risk-reduction-2019>.

Wijayanti, P, Zhu, X, Hellegers, P, Budiyono, Y, van Ierland, EC 2017, ‘Estimation of river flood damages in Jakarta, Indonesia’, *Natural Hazards*, no. 86, pp. 1059-1079.

World Bank 2008, ‘Project information document (PID): concept stage – Jakarta Urgent Flood Mitigation Project (Report No. AB4043)’, *East Asia and Pacific Region*. Government of Indonesia.

Appendix

Section A: Tables and Figures

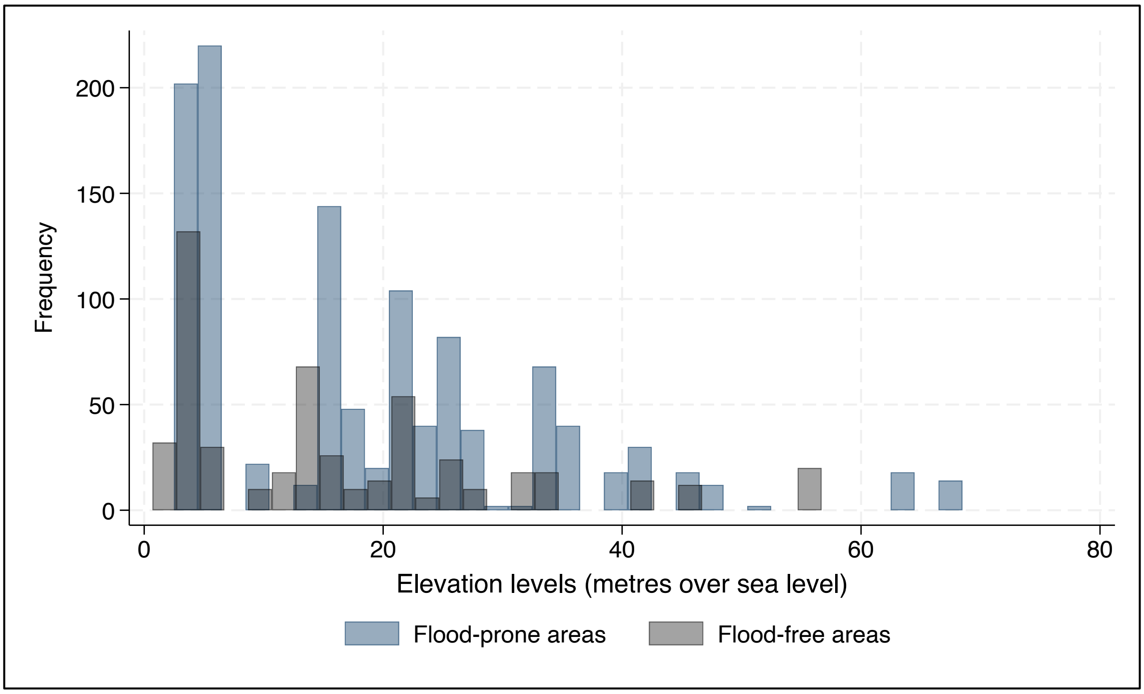


Figure 2.A.1 Elevation difference between flood-prone and flood-free areas

Source: The Impact of Jakarta Floods Survey (2018).

Table 2.A.1 Distribution of the degree of risk aversion

Degree of risk aversion		Distribution of sample when: Hypothetical payoffs		Real payoffs		Tendency to move to riskier option
Extreme	IDR 20,000 - IDR 20,000	256	31%	240	29%	-6.3%
Severe	IDR 16,000 IDR 25,000	160	19%	141	17%	
Intermediate	IDR 12,000 - IDR 30,000	166	20%	155	19%	
Moderate	IDR 8,000 - IDR 35,000	80	10%	79	9%	
Slight-to-neutral	IDR 4,000 - IDR 40,000	86	10%	117	14%	
Neutral-to-negative	IDR 0 - IDR 50,000	88	11%	104	12%	27.0%
Total		**836**	**100%**	**836**	**100%**	

Source: The Impact of Jakarta Floods Survey (2018).

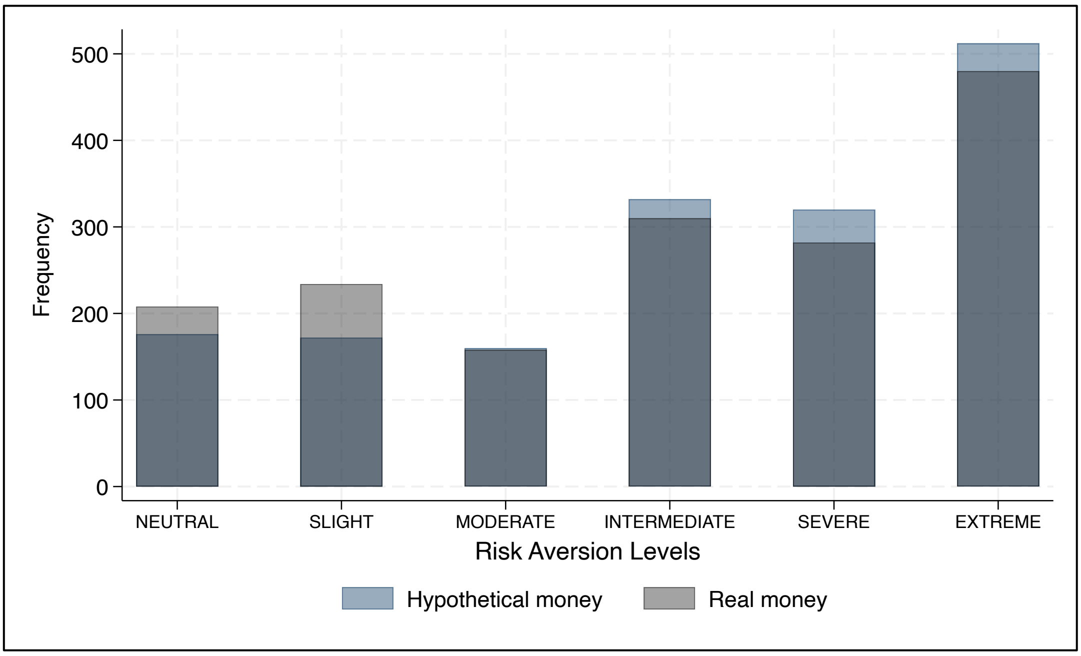


Figure 2.A.2 Risk aversion difference between hypothetical and monetary games

Source: The Impact of Jakarta Floods Survey (2018).

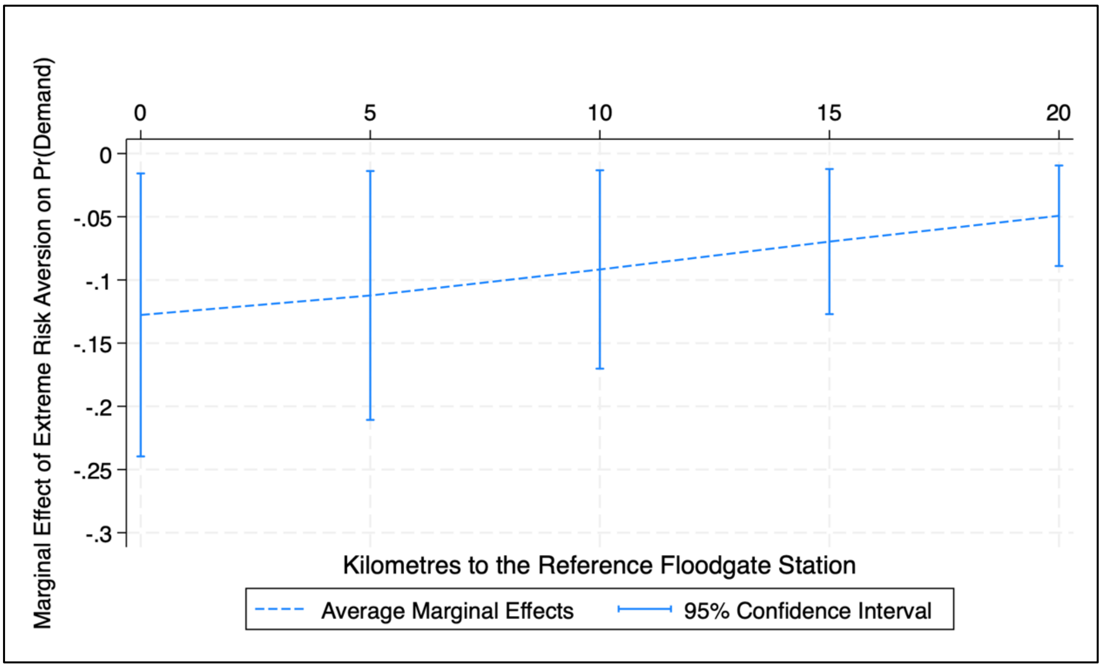


Figure 2.A.3 Extreme risk aversion of demand and distance to reference floodgate station

Source: The Impact of Jakarta Floods Survey (2018).

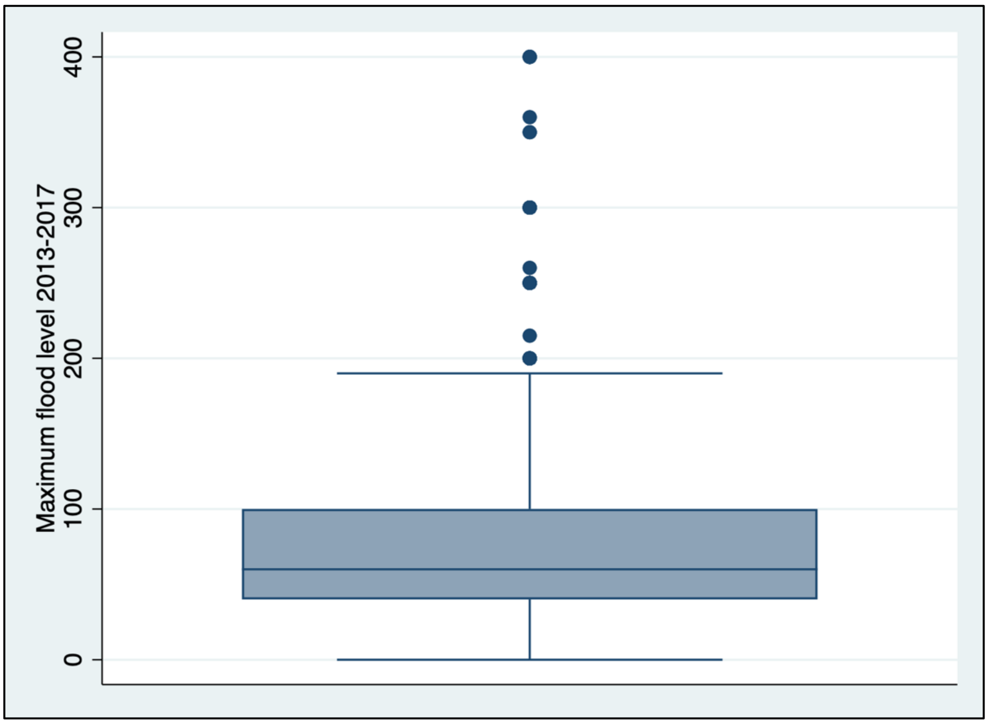


Figure 2.A.4 Outliers regarding the urban villages with flood levels above 2 metres

Source: The Impact of Jakarta Floods Survey (2018).

Table 2.A.2 Index-based flood insurance demand among households with real monetary payoffs

	Dependent variable: Insurance demand							
Model: Probit	(1) All sample		(2) 1st question		(3) 2nd question		(4) All sample	
Distance to floodgate (km)	-0.014	***	-0.016	***	-0.011	**	-0.014	***
	(0.003)		(0.004)		(0.004)		(0.003)	
Premium (in logs)	-0.133	***	-0.153	***	-0.106	***		
	(0.009)		(0.009)		(0.015)			
Extreme risk-aversion (1,0)	0.021		0.017		0.023		-0.028	
	(0.019)		(0.026)		(0.028)		(0.043)	
Distance to floodgate (km) x Extreme risk-aversion (1,0)							0.005	
							(0.004)	
Distance to nearest river (in logs)	Yes		Yes		Yes		Yes	
Region FE	Yes		Yes		Yes		Yes	
Flood intensity FE	Yes		Yes		Yes		Yes	
Covariates	Yes		Yes		Yes		Yes	
Number of observations	1,662		831		828		1,662	

Note: This table presents probit regression estimations of household-level insurance uptake on distance to the reference floodgate station, premium and extreme risk-aversion (with real monetary payoffs). All sample corresponds to double-bounded dichotomous choice questions, except for columns (2) and (3). The "Far from and close to floodgate station'' sample only includes households at a perpendicular distance from their reference floodgate station below 5 km (''close'') or over 12 km (''far''). In columns (4), (5) and (6) probit specifications include the interaction term. All specifications include distance to nearest river, covariates, as well as region and flood intensity fixed effects. Standard errors are bootstrapped (1,000 replications) and reported in parentheses. Asterisk (*), double asterisk (**), and triple asterisk (***) denote variables significant at 10 percent, 5 percent, and 1 percent, respectively.
Source: The Impact of Jakarta Floods Survey (2018).

Table 2.A.3 Premium rates and demand for index-based flood insurance

Model: Probit	Dependent variable: Insurance demand					
	(1) All sample		(2) 1st question		(3) 2nd question	
Premium rates						
IDR 10,000	0.498	***				
	(0.084)					
IDR 20,000	0.590	***			0.634	***
	(0.090)				(0.095)	
IDR 25,000	0.066				0.077	
	(0.083)				(0.078)	
IDR 50,000	0.115		-0.327	***	0.062	
	(0.077)		(0.054)		(0.071)	
IDR 100,000	-0.084		-0.589	***	-0.050	
	(0.078)		(0.052)		(0.068)	
IDR 200,000	-0.113		-0.618	***	-0.074	
	(0.076)		(0.050)		(0.075)	
IDR 400,000	-0.010				0.048	
	(0.120)				(0.131)	
IDR 500,000	-0.144	*	-0.649	***		
	(0.077)		(0.048)			
IDR 1'000,000	0.287	*			0.333	*
	(0.165)				(0.177)	
Distance to floodgate station (km)	-0.011	***	-0.015	***	-0.007	*
	(0.003)		(0.004)		(0.004)	
Extreme risk aversion (1,0)	-0.040	*	-0.031		-0.046	
	(0.022)		(0.027)		(0.030)	
Distance to nearest river (in logs)	Yes		Yes		Yes	
Region FE	Yes		Yes		Yes	
Flood intensity FE	Yes		Yes		Yes	
Covariates	Yes		Yes		Yes	
Number of observations	1,553		831		719	

Note: This table displays probit regressions of index-based flood insurance demand on distance to the reference floodgate station, premiums, and extreme risk aversion. Premium rates come from the five price levels and their follow-up (discounted or increased) price under a double-bounded dichotomous choice. All price levels are dummy variables, except for IDR 5,000 (USD 0.35) which is omitted in the table. Average marginal effects are reported. All sample corresponds to double-bounded dichotomous choice questions in column (1), and the first and follow-up questions in columns (2) and (3), respectively. Standard errors are bootstrapped (1,000 replications) and reported in parentheses. Asterisk (*), double asterisk (**), and triple asterisk (***) denote variables significant at 10 percent, 5 percent, and 1 percent, respectively.
Source: The impact of Jakarta floods survey 2018.

Table 2.A.4 Estimation of Mean WTP (MWTP)

Model: OLS	Dependent variable: insurance demand
	(1)
Premium (in logs)	-0.143 ***
	(0.009)
Constant	0.820
	(0.080)
Distance to near river (in logs)	Yes
Region FE	Yes
Flood intensity FE	Yes
Covariates	No
Number of observations	1,672
Index insurance premium (monthly)	IDR 167,000 (USD 11.53)
MWTP	IDR 8,299.65 (USD 0.57)

Note: This table presents an OLS regression of index-based flood insurance demand on the premium, based on a double-bounded dichotomous choice question. The mean willingness-to-pay (WTP) is calculated using the following formula: $\left(-\frac{\log(1+\exp(0.820))}{-0.143}\right) * 1000$. Standard errors are bootstrapped (1,000 replications) and reported in parentheses. Asterisk (*), double asterisk (**), and triple asterisk (***) denote variables significant at 10 percent, 5 percent, and 1 percent, respectively.
Source: The impact of Jakarta floods survey 2018.

Section B: Estimation of flood shocks[1] in Jakarta

The index-based insurance literature usually employs weather-related shocks as a source of exogenous variation to identify effects associated with insurance coverage. Similarly, in this study I harness the substantial changes in flood levels within a particular spatial unit over a five-year period (2013–2017) in Jakarta, assuming these changes have positive, direct effects on households' willingness to purchase hypothetical index-based flood insurance. If changes in flood levels are not included as a control variable, there may be potential bias in the probit estimates. Such endogeneity concern relies on the idea that some households may be sensitive to flood level variations (i.e., the elderly and pregnant women). Therefore, I measure the variation in flood levels deviations registered at *kecamatan* (subdistrict level) during 5 rainy season months of the 2013–2017 period.

The rainy season brings very intensive rainfall typically between the months of October and April (Texier 2008) and increases the number of flood events in Jakarta; however, most occur from November to March. Figure 2.A.5 shows the number of villages affected by flooding within this period.

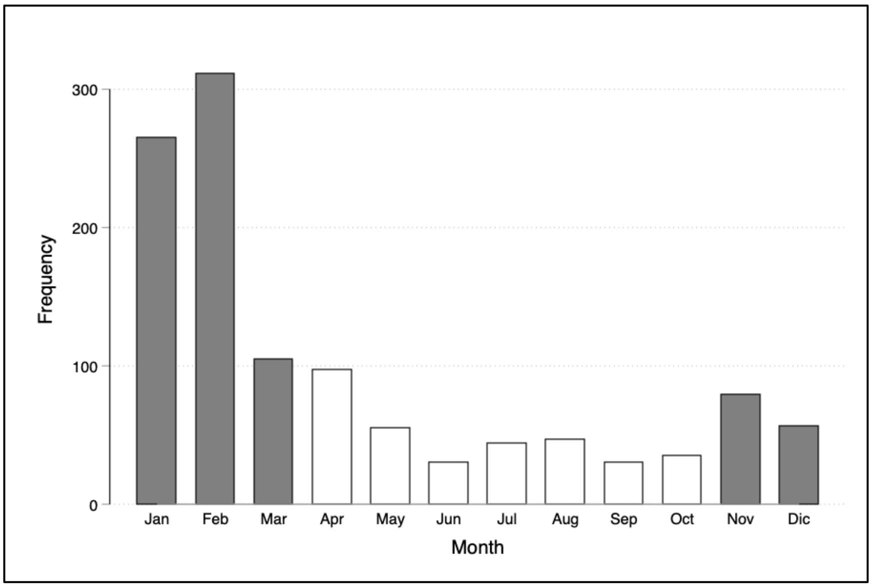


Figure 2.A.5 Village flood occurrences during 2013-2017

Source: The impact of Jakarta floods survey 2018

[1] Flooding data obtained from Jakarta Regional Disaster Management Agency – BPBD. Information was collected in the 2013-2017 period as follows: i) identify the flood areas at urban village level using GPS devices; ii) at flood locations, depth marks on home walls – water limits of flood depth – were registered as historical data; and iii) ask households for date of previous flooding and time of flooding to recede in their respective neighbourhood. The data includes 204 villages out of a total of 261.

The estimation of rainy season flood shocks in 2013-2017 follows the approach of Karadja and Prawitz (2017), expressing shocks measured in relation to the local long-run weather in a particular month. First, for each month m, year t and location l, I calculate the deviation between the actual and the long-run average flood level registered in that month:

$$deviation\ (Flood\ Level)_{l,m,t} = Flood\ Level_{l,m,t} - \overline{Flood\ Level}_{l,m}$$

where a flood level is the height of flood event occurring in the $kecamatan$. A flood shock is then defined as a binary variable:

$$shock_{l,m,t} \equiv I[deviation\ (Flood\ Level)_{l,m,t} > sd(Flood\ Level_{l,m})]$$

where $shock_{l,m,t}$ is equal to one if the $kecamatan\ l$ experienced a positive flood shock in a month m of year t. The $kecamatan$'s long-run standard deviation of flood level in each month over the 2013-2017 period is denoted by $sd(Flood\ Level_{l,m})$[2]. Lastly, I sum the number of shocks over the rainy season for each $kecamatan$ between 2013 and 2017. The frequency distribution of flood shocks during 2013-2017 is displayed in Figure 2.A.6, with the median *kecamatan* experiencing four flood shocks. The observations of flood shocks are matched with survey data at subdistrict level.

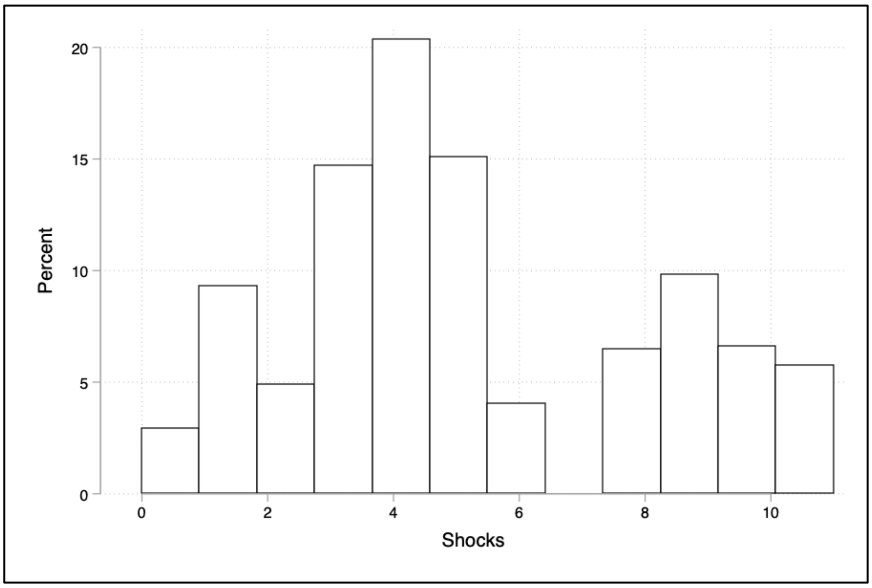


Figure 2.A.6 Frequency distribution of flood shocks 2013-2017

Source: The impact of Jakarta floods survey 2018

[2] In order to fill in the missing values (24 observations out of a total of 816), I use that of the nearest kecamatan to complete the long-run standard deviation in a particular month.

Chapter 3
Under the water: flood impacts and economic dynamics in northern Peru

3.1 Introduction

Climate change increases the frequency and intensity of extreme weather-related events globally. These events are sudden and sudden and unexpected, resulting from deviations in climate patterns from normal environmental conditions across large regions. Among the most common and damaging natural hazards, floods have caused more economic and property damage, and loss of life, than any other natural hazard in the 21st century (CRED, 2018). The "El Niño" phenomenon exacerbates flood risks and the likelihood of long-lasting and devastating effects on people's livelihoods, particularly in developing countries where vulnerability is often linked to poverty. One of the countries most affected is Peru (Tyndall Centre, 2004), with approximately 1.51 million people living in inundation zones or areas susceptible to floods (Smith et al., 2021).

In early 2017, Peru experienced severe floods due to the coastal El Niño – a variant of the "El Niño" phenomenon, causing deaths and infrastructure damage (OCHA, 2017). The National Meteorology and Hydrology Service reported 107 fatalities (Sardon et al., 2022), while the Ministry of Economy and Finance noted a 1.5 percentage points[1] decline in Peru's Gross Domestic Product (GDP), estimated at US$ 3.1 billion as at 28 March 2017 (Leon et al., 2017). The floods affected the northern region, including the departments of Cajamarca, La Libertad, Lambayeque, Piura, and Tumbes. Residents faced recurring floods triggered by heavy rainfall in February and March of 2017, with the most extreme episodes occurring between 26 and 27 March 2017. Subsequent events in 2018 and 2019 sporadically affected a few districts within the same region, but with lesser impact and damage compared to the 2017 floods.

An investigation into the impact of extreme natural events on households in northern Peru is therefore crucial. It aids in understanding household disaster recovery and assesses the risk of

[1] According to the Central Bank's Annual Report Memory 2017, there were no transitory shocks or other economic disruptions apart from the 2017 coastal El Niño, which negatively impacted the entire economy of the northern region of Peru. This report can be accessed from the following link: https://www.bcrp.gob.pe/docs/Publicaciones/Memoria/2017/memoria-bcrp-2017.pdf

poverty trap if there is limited external support. This study examines the economic effects of flood shocks on household welfare and poverty status by using five–year panel data from 2015 to 2019 and employing a diff-in-diff event study model. The identification strategy relies on the causal effect of flood variation on household income and expenditure per capita, and poverty status in areas affected by this natural hazard relative to households in unaffected areas. Using a novel damage index as a proxy for the local impact of floods, treated households are defined as those located in affected districts, and control households are those that lived in unaffected districts – discrete specification. Additionally, I use the index as a continuous measure of treatment as treated households were exposed to different levels of flood damage. I then compare treatment and control groups to estimate the effects of such disaster on household welfare (income and expenditure) and poverty status at the district level.

I construct the damage index as a measure of local economic impact for districts affected by floods, using several publicly available remote sensing such as watershed shapefiles, Digital Elevation Model, Land Use/Land Cover, and Digital Soil Map of the World rasters, weather data, and night-time lights from the Visible Infrared Imaging Radiometer Suite (VIIRS). The estimates of the associated damage may be useful for policy making prior to the realisation of a potential extreme weather event such as flooding. I also consider the possibility that households have relocated to areas with lower or non-flood risk after the occurrence of flooding, which may lead to inaccurate estimations of flood impacts. To address these endogeneity concerns due to the sorting of flood risk areas[2], I conduct propensity score matching in the pre-estimation step to create counterfactual observations of households in non-flooded areas. This ensures a more accurate analysis of the relationship between floods and household welfare in northern Peru.

The econometric results suggest that, in response to an average flood shock, household income and expenditure fall by 29% and 26.5%, respectively. I also find that poverty in urban areas increases by 41% because of job loss and deterioration of local economic activity. I observe signs of recovery in household income and expenditure but estimates are lower than those in the year prior to the flood occurrence.

---

[2] This is noted by Bakkensen and Ma, 2020 in the context of individuals sorting across levels of flood risk (low and high) by race, ethnicity, and income for South Florida, USA.

The contribution to the literature is threefold. First, to the best of my knowledge, this study is one of the few research efforts to propose an index for measuring economic damage and estimating the causal impacts of the 2017 floods on districts within the northern region of Peru. The index not only serves as a proxy for local economic impacts but also ranks districts from most to least economically affected. This approach can provide central and local governments, emergency services, and aid workers with a practical tool to estimate the damages and intensity of the disaster—either ex ante or ex post—and allocate resources to areas most in need. Furthermore, previous empirical studies have usually relied on precipitation data as a proxy for floods, rather than capturing this weather-related event at the surface level.

Second, recent contributions in econometrics have highlighted the pitfalls of time-within-fixed effects models in study settings with multiple groups and time periods, variation in treatment timing, and a continuous treatment (Callaway and Sant'Anna, 2021; de Chaisemartin and D'Haultfœuille, 2023; Goodman-Bacon, 2021; Sun and Abraham, 2021). In the study design, due to the nature of the variable of interest, I evaluate the causal effect of flood events on household welfare using the heterogeneity robust estimator proposed by de Chaisemartin and D'Haultfœuille (2023). This approach strictly compares districts switching to treatment in a time period with groups not-yet-treated during the same period and specifically considers dynamic effects.

Third, I provide additional evidence on the short- to medium-term impacts of floods on household income and expenditure per capita, poverty status, number of income earners, wage income, savings and disaster relief (food and clothing) by using five-year household survey and remote sensing data, as well as Peruvian government reports of the districts' exposure to 2017 flood impacts.

The paper proceeds as follows. Section 3.2 reviews relevant literature, focusing on flood events that have occurred in developing countries and their impacts on household welfare. In Section 3.3, I describe the coastal El Niño floods that occurred in northern Peru in 2017. Section 3.4 explains the empirical strategy, including the classification of districts as treated and control groups, and the empirical modelling. Section 3.5 provides information on the datasets and the construction of the damage index, while Section 3.6 presents the pre-estimation process to address potential endogeneity resulting from sorting across households.

Sections 3.7 presents the main results and extended analysis. Lastly, Sections 3.8 and 3.9 provide policy recommendation and concluding statements, respectively.

3.2 Related literature

The effect of weather-related shocks has been documented in several studies. However, there is little assessment of the impact of sudden-onset events on household incomes and expenditures in middle-income countries. The evaluation that does exist is focused either on households in high-income countries (e.g., earthquakes in Japan, and hurricanes in the United States) or on rural households in low-income countries (e.g., droughts in sub-Saharan Africa, and hurricanes in Central America).

The literature on flood impacts finds that such events negatively affect the households' welfare, economic preferences, subjective expectations, and behavioural choices. In Thailand, Noy et al. (2020) find that the 2011 floods caused significant income and expenditure losses among affected households, with decline in business and wage earnings driving these outcomes. Interestingly, unaffected households in flood areas ("spillover" households) experienced income losses as much as the directly impacted households. Floods also led to higher housing-related expenses but lower spending on luxuries, particularly among wealthier households. Chantarat et al. (2016) report that flood-affected households tend to be more risk averse, impatient, and altruistic, with asset-poor farming households more vulnerable than their better-off counterparts. Moreover, floods increase households' expectations of experiencing severe flooding in the future.

In Bangladesh, Del Ninno et al. (2001) examine the 1998 flood's effects on rural household welfare, finding that more than half of those affected suffered losses in assets, employment, and agricultural workdays. They highlight coping mechanisms such as reducing expenditure, selling assets, and —most commonly— borrowing. Government relief, including food and cash transfers, was targeted to flood-affected and poor households but fell short of meeting all needs. Similarly, Mueller and Quisumbing (2011) use a five-year household panel data to assess both short- and long-term flood impacts on wages. They show short-term decline in wages of salaried workers in the non-agricultural sector (8.4%–13.8%) and agricultural sector (34.3%–45.6%). Agricultural workers who shifted to non-agricultural jobs as a coping mechanism experienced smaller short-term wage reductions.

Another branch of literature explores the health consequences of floods. Yonson et al. (2018) estimate the probability of urban households in the Philippines experiencing diseases between 2011 and 2014. They find that higher floodwaters are linked to increased cases of bronchitis, influenza, and leptospirosis, while longer flood duration is associated with bronchitis, influenza, and typhoid fever. In the same way, using state-level panel data from 17 Indian states for 1995–2011, Parida et al. (2018) analyse the relationship between floods and farmer suicides, finding no direct link.

Maccini and Yang (2009) investigate the effects of excessive rainfall in Indonesia on various adult outcomes. They argue that rainfall shocks occurring around the time of birth—well before schooling age—can have lasting effects on health, education, and socioeconomic status. These operate through early-life health and schooling pathways. Higher early-life rainfall is associated with better health indicators, taller stature, higher educational attainment and socioeconomic status, particularly among women. The authors attribute these positive effects to rainfall's boost to agricultural yields and household income.

Research on disasters and poverty highlights disproportionate effects on poorer households. Using a meta-regression of household-level studies, Karim and Noy (2016a, 2016b) find that low-income households tend to smooth food consumption by reducing non-food spending, particularly on health and education, suggesting long-term adverse consequences and thereby perpetuating poverty. Rodriguez-Oreggia et al. (2013) also show that disasters generally increase poverty, with floods and droughts having more severe effects than frost or heavy rainfall. Janzen and Carter (2013) bridge the literature on post-disaster poverty traps, asset dynamics, and microinsurance. Studying drought-affected households in northern Kenya, they use instrumental variables to account for selection bias and find that insurance payouts reduced the likelihood of asset depletion by 22–36 percentage points. Households with assets above a certain threshold are more likely to smooth consumption, while those below it display asset-smoothing behaviour.

External meteorological measures, such as rainfall data (Lertamphainont and Sparrow, 2016; Kocornik-Mina et al., 2020) or inundation maps, including satellite images (e.g., Poaponsakorn et al., 2015; Noy et al., 2020), are often combined with household surveys to assess flood impact. For instance, Noy et al. (2020) merge household self-reported shocks with satellite images, while Chantarat et al. (2016) use a flood inundation map to distinguish heavily affected villages from nearby, less affected ones. Mueller and Quisumbing (2011) construct a flood shock variable based on deviations in flood depth from normal conditions, and Yonson et al. (2018) measure three aspects of flooding—height in meters, duration in hours, and exposure. Parida et al. (2018) identify flood–affected areas using flood maps. Beyond floods, Maccini and Yang (2009) estimate exogenous variation in rainfall as the percentage deviation of birth-year rainfall from the long-run average in a given municipality.

However, these approaches are often constrained by limited frequency and measurement imprecision (Noy et al., 2020). Moreover, as Guiteras et al. (2015) note, flooding in a specific area can be influenced by a broad and complex set of hydrological conditions[3], making it crucial to account for these conditions when estimating economic impacts. To address this, the present analysis constructs a district-level damage index following the approach of Skoufias et al. (2020). Damage is defined as the local economic impact of flooding within a district, estimated by using open-source remote sensing data typically applied in natural hazard modelling, combined with night-time lights as a proxy for economic activity. This method calculates river discharge values within a basin, providing a more accurate representation of flood impacts in northern Peru. The details of the index construction are outlined in Section 3.5.4.

3.3 The 2017 'coastal' El Niño in Peru

The coastal El Niño (El Niño costero in Spanish) is one of the El Niño variants and occurs specifically in northern Peru. Unlike the traditional El Niño/Southern Oscillation – ENSO[4], that involves warming of ocean waters in the central and eastern Tropical Pacific (Niño3.4 region), the coastal El Niño events are localised and characterised by warmer-than-average sea surface temperatures (SSTs) along the coast of the departments of Tumbes, Piura, Lambayeque, and La Libertad (Niño1+2 region). These events predominantly occur between February and April when SSTs are typically at their warmest. A notable feature of coastal El Niño is the rapid and abrupt increase in SSTs, often rising by 7–9°C over a period of 1-2 weeks (Takahashi et al., 2017). This signals the beginning of heavy rainfall, usually occurring when surface temperatures exceed approximately 28˚C (Barnston, 2017).

In February to March 2017, the floods in northern Peru came as a surprise (Venkateswaran et al., 2017). Unfortunately, El Niño researchers, including those in the United States and Peru,

[3] Their findings show a positive but modest correlation (0.09) between monthly rainfall and flood extent measured at the district level in Bangladesh over the 2002 to 2011 period.

[4] Peruvian fishermen were the first to coin the term "El Niño" in the late 1800s after noting unusual oceanic warmth near their coast.

the United States National Oceanic and Atmospheric Administration (NOAA), and the regional and Peruvian meteorological agencies did not expect this extreme weather shock despite a basin–wide monitoring system across the Pacific[5]. Thus, government agencies, socioeconomic sectors, public health officials, and citizens were not forewarned of the potential extreme weather and environmental-related hazards such as floods, flash floods and landslides. In addition, there was confusion over whether conditions in 2017 represented an El Niño phenomenon – as experienced in 1982 to 1983 and 1997 to 1998 – that provoked a hydrometeorological debate and stifled decision making (Ramirez et al., 2017).

Based on the January 2017 update blog from the NOAA ENSO alert system, La Niña conditions (a cooling pattern in the eastern Tropical Pacific) were anticipated for December 2016, after the monthly SST Niño 3.4 Index values dropped below the threshold of –0.5°C (Becker, 2017). However, on 31 January 2017, the SST abruptly increased reaching values above 26°C at various points along the northern coast in the Niño1+2 region (Multisectoral Committee for the Study of El Niño – ENFEN, 2017). These SST anomalies informed the Peruvian government that they were dealing with El Niño-like rainfall conditions. Only a few days later, on 6 February 2017, rainfall brought floods to the departments of Tumbes, Piura, Lambayeque, La Libertad, and Cajamarca.

On 23 February 2017, the government declared a state of emergency in many affected areas in northern Peru, and international aid efforts were mobilised to assist those impacted by floods. The worst flooding occurred between 26 and 27 March 2017 when heavy rainfall combined with already high river flows. These adverse effects associated with the SST anomalies can be seen in Figure 3.1, which displays the significant difference between the long-term average SST in March from 1980 to 2010 and the SST registered in March 2017 (highlighted in dark red) along the northern coast of Peru. By May 2017, the coastal El Niño finally subsided, leaving behind losses estimated at US$ 3.1 billion, as at 28 March 2017 (Leon et al., 2017), with close to 1.1 million people affected and approximately 46.7 thousand destroyed houses, as at 10 May 2017 (OCHA, 2017).

---

[5] The Niño 3.4 region is important for predicting and monitoring basin-wide El Niños; however, it may not always be the best area for observing other types of El Niño events that have extreme weather impacts concentrated along the coasts of northern Peru and southern Ecuador.

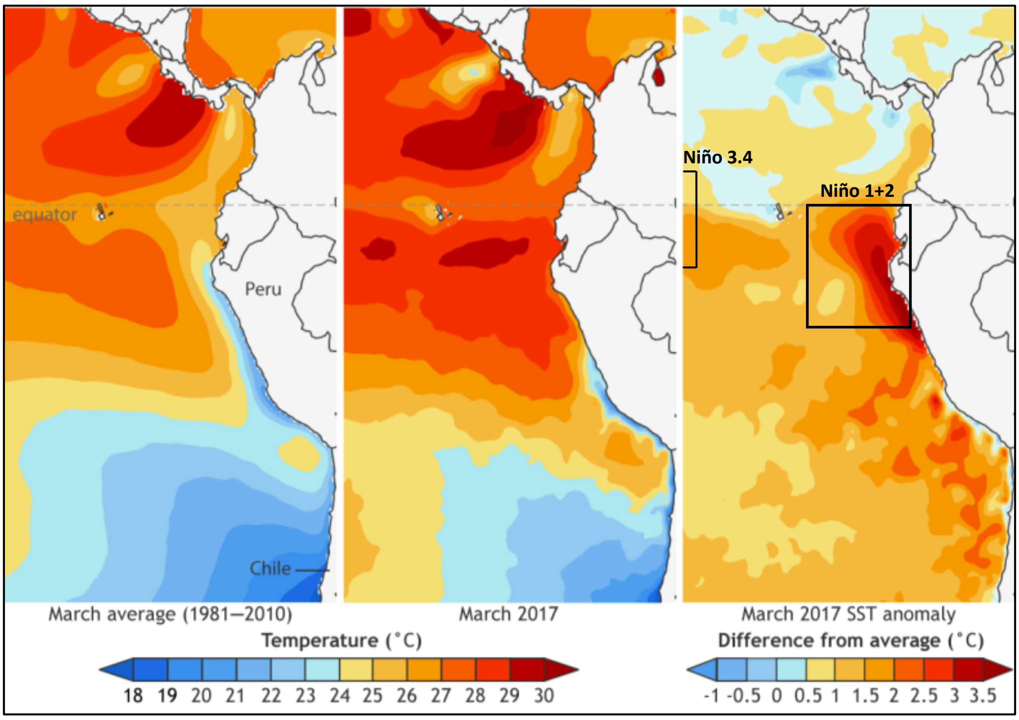


Figure 3.1 SST anomaly in March 2017 along the northern coast of Peru

The left panel displays the long-run average of sea surface temperatures (SST) in the eastern Tropical Pacific for March from 1981 to 2010. In the middle panel, SST for March 2017 is shown. The right panel illustrates the difference between the left and middle panels, indicating the SST anomaly for March 2017 along the northern coast of Peru.
Source: NOAA Climate.gov – April 2017 ENSO Update.

3.4 Empirical strategy

Treatment variable

The treatment variable represents the local economic impact (i.e., at the district level) resulting from floods. It is defined as a binary variable that takes a value of one if the district was affected in 2017, switches to zero if the district became unaffected, and reverts to one if it was affected again, or remains one if the district was continuously affected between 2018 and 2019 (see Figure 3.A.1 in the Appendix). For districts that were never affected, the treatment variable remains zero throughout. The binary treatment is based on an index associated with the damage to buildings and infrastructure calculated using remote sensing data at the district level[6]. Due to floods being very localised, the damage index appears to determine the financial and property losses suffered by affected districts relative to unaffected districts in northern Peru.

[6] The index includes the flood and night-time light intensities to take into account the local physical characteristics of this natural hazard, and the local economic activity exposed to it.

The decline in household welfare and the increase in poverty were plausibly exogenously determined by flood events. However, this approach may be confounded by residential sorting. As households relocate across areas with varying flood risk within a district, the magnitude of the flood impacts may become biased. It is problematic to take the observations of these affected households; therefore, in Section 3.6.1, I conduct the Propensity Score Matching (PSM) technique and drop these observations from the analysis. The newly matched data, adjusted with both PSM and household weights, will be used for the rest of the study.

Definition of treatment and control groups

To efficiently capture variation in the impact of flooding, I define two groups: one with affected districts, and one with unaffected districts. First, I define a measure of treatment, which is the damage index used as a proxy for the local economic impact in flooded districts. Second, the group with impacted districts is identified if the damage index is greater than zero (**treatment**) and the group with free-impacted districts if the damage index is equal to zero (**control**).

$$FD_{d,T} = \begin{cases} 1 \;\; if \;\; damage\ index > 0 \\ 0 \;\; if \;\; damage\ index = 0 \end{cases}$$

Estimating treatment effects

In this study, the key identification assumption is that the 2017 coastal El Niño floods occur unexpectedly regardless of household behaviour. This allows us to compare households in affected districts to those in unaffected districts. I causally estimate the economic impact of these weather-related events on their income and expenditure per capita, and poverty status using a standard diff-in-diff/event study model. The estimating equation is written as:

$$y_{h,T} = \beta_0 + \delta_{-2}.FD_{d,T}^{-2} + \sum_{e=0}^{L=2} \boldsymbol{\tau_e}.FD_{d,T}^{e} + \mu_d + \lambda_T + X_{h,T}.\gamma + \varepsilon_{h,T} \qquad (1)$$

Subscripts h, T and d denote the household, year, and district respectively; $y_{h,t}$ is the household welfare and poverty status outcomes; $FD_{d,T}$ is a dummy variable indicating whether the district was affected by floods at year T (=1) or not (=0); μ_d and λ_T are the district and year fixed effects, respectively; $X_{h,T}$ is a vector of covariates, including dwelling information, such as dummy indicators for access to sewage, water, electricity, and a concrete roof; household head characteristics, such as dummy variables for female status and public health insurance coverage, years of education, and age in years; household-level variables, such as household size, the number of income earners, and a dummy variable indicating whether the household owns their own house or not; and geographic characteristics, such as distance to the near river, terrain slope and altitude. For household welfare and poverty status outcomes, $e = -2$ specifies the year 2015 before the 2017 coastal El Niño floods, while $e = -1$ is the (dropped) base year against which all estimated effects are compared; $e = 0$ signifies the year of the 2017 coastal El Niño floods; and values of $e > 0$ and $\leq L$ indicate the years after the 2017 coastal El Niño floods. β_0, δ_{-2}, and τ_e are coefficients to be estimated.

In the standard diff-in-diff/event study model defined in equation (1), estimated δ_{-2} indicates pre-treatment impacts and estimated $\boldsymbol{\tau_e}$ denote average treatment effects. The latter can be taken as causal under the supposition that both the "parallel trends" and "no anticipation" assumptions are satisfied. The parallel trends assumption implies that the outcome trend of the treated group would have evolved in parallel to the outcome trend of the control group, in both pre- and post-treatment periods, given no treatment. The no anticipation assumption requires that no treatment effects exist before treatment begins. A more detailed discussion of these two assumptions is found in Section 3.6.2.

Since the panel data include only one lead year (2015), I cannot directly test the parallel trends assumption[7]. To address this limitation in this study, I ensure that the data are balanced in household characteristics between the treated and control groups in the pre-intervention period (2015–2016). Using propensity score matching (Section 3.6), I achieve more balanced conditions and then run the DiD event study regression on the matched sample only, as the parallel trends assumption is more likely to hold when covariates are balanced prior to treatment. This technique creates counterfactuals for the treated group and excludes observations without a counterfactual pair (off-support) from the sample. If the estimated coefficient on the lead term, δ_{-2}, is statistically insignificant, this is generally considered supportive evidence for the parallel trends assumption (Sun and Abraham, 2021).

---

[7] Roth (2022) and Rambachan and Roth (2023) developed two test procedures to support the claim that the parallel trends assumption holds. However, it is not possible to apply these procedures in this study due to the availability of only one year of data (i.e., year 2015) before the flood shocks in 2017.

Equation 1 defines the conventional diff-in-diff/event study model. One challenge of this framework is the issue of the "forbidden comparisons" or "negative weights." This approach estimates treatment effects by comparing treated units to control units, which may include "never treated" units, "not-yet treated" units, and "already treated" units. While "never treated" and "not-yet treated" are valid controls, "already treated" controls are potentially problematic because they may receive negative weights in estimation. These negative weights are likely to bias the estimated effects in both the pre- and post-treatment periods (Goodman-Bacon, 2021; Callaway and Sant'Anna, 2021; Sun and Abraham, 2021).

The identification strategy in this study assumes that the treatment variable is quasi-random after controlling for distance of the surveyed house to the nearest river. In the Appendix (see Tables 3.A.1 and 3.A.2)[8], I provide empirical evidence supporting this assumption. Note that the quasi-random treatment variable assumes that it is orthogonal to both potential outcome levels and trends, whereas the parallel trends assumption relies on the independence of potential outcomes trends only (Roth and Sant'Anna, 2023). While the treatment variable is plausibly exogenous, the data contain only one pre-treatment year and lack time trends. As a result, it cannot be argued that the quasi-random treatment assumption is stronger than the parallel trends assumption. Therefore, I include covariates in the estimation to help ensure that the parallel trends assumption holds, following the recommendations of Roth et al., (2023) who advise conditioning on covariates particularly when concerns exist about the validity of your parallel trends assumption without including covariates. A more detailed discussion of parallel trends and no anticipation assumptions is found in Section 3.6.2.

In addition, negative weights may still be problematic, however, even when treatment is plausibly exogenous and, as a result, the standard diff-in-diff/event study estimation remains ill-advised (Roth and Sant'Anna, 2023). To circumvent the forbidden comparisons and negative weights problems, as well as accommodate the treatment that switches on and off over time, I employ an estimator developed by de Chaisemartin and D'Haultfuille (2023).

[8] I take two distinct approaches. First, following Lewis (2017) and Lewis and Dong (2025), I regress the binary switching treatment and continuous treatment on pre-determined outcomes and covariates using data from 2015 to 2017. As shown in Table 3.A.1, none of the predetermined outcomes or covariates explains the first year of flood shocks at standard levels of statistical significance (i.e., five percent), except for household savings (for binary switching treatment) and disaster relief (for continuous treatment) at 1% and 5% significance level, respectively. Second, adapting the approach of Lewis and Dong (2025) to the study, I repeat this procedure on the subsample of "never treated" districts, using data from the year before the coastal El Niño floods (2016) through 2019. Table 3.A.2 shows that none of the outcomes or covariates explains the floods at any levels of statistical significance. These results support the assumption that the flooding was quasi-random.

This modern event study estimator starts by aggregating units into cohorts of "switchers" and "non-switchers" based on the year in which they received treatment. Note that the non-switchers include two groups: (i) those whose treatment status stays the same between periods; and (ii) those with no treatment status between periods. The estimator then compares the evolution in mean outcomes among cohorts that switched into treatment between periods $t-1$ to t to the evolution in mean outcomes among groups that were untreated in both periods $t-1$ to t[9]. It then averages these diff-in-diff estimands across all groups to derive the average treatment effect on the treated (ATT)[10]. Therefore, the estimator is robust to heterogeneous effects across groups of all treatments, and ensures that it is robust to heterogeneous effects over time of all treatments. I test the robustness of the main results in this paper using the Dynamic Two–Way Fixed Effects (TWFE) approach, which I discuss more in detail in Section 3.7.2.

This paper uses the Stata package 'did_multiplegt' developed by de Chaisemartin and D'Haultfuille (2023)[11] to estimate the main treatment effects. The analysis includes control variables, 200 bootstrap replications, and clusters standard errors at the district level. Since the estimator does not restrict treatment to binary form (i.e., it allows for continuous treatment), I use the flood damage index based on remote sensing data as a proxy for the local economic impact to demonstrate that the results are consistent with those from the switching treatment dummy specification.

3.5 Data

The analysis is based on a district-level dataset covering demographic, socio-economic, geographic, and weather variables over the period 2015 to 2019 from multiple sources. In addition, I utilise Peru's government reports on flood events registered between 2017 and 2019 to identify the districts typically affected. This also includes information on government

9 Switchers can only switch from untreated to treated between periods $t-1$ to t, and this may occur at different points in time. This likely aligns with the definition of a staggered design; however, it extends the definition of a staggered design to include non-binary treatments as it is in this study.

10 The Average Treatment Effect on the Treated (ATT) is accurately identified when additionally, the parallel trends and no anticipation assumptions already hold.

11 This study uses the older version of the did_multiplegt Stata package instead of the newer did_multiplegt_dyn version, as the latter does not account for some control variables in groups (therefore, less observations used to estimate the Average Treatment on the Treated) where the baseline treatment is equal to 0. This issue may arise because some control variables in the dataset lack sufficient variation over time and across groups, particularly those related to geographic characteristics.

support, international aid, infrastructure affected by floods, and the number of affected households and people displacement.

3.5.1 Household data

The panel data comes from the Peruvian National Household Survey (ENAHO) collated by the National Institute of Statistics and Informatics of Peru. A total of 308 households has been tracking over five years, from 2015 to 2019, across 57 districts in northern Peru. The survey includes detailed information on socioeconomic characteristics over time for each household, as well as demographic information such as gender, age, marital status, and educational attainment.

An advantage of using ENAHO is that it provides information on household coordinates (longitude and latitude), and individual places of residence at the time of the survey[12]. This is contrary to previous studies that use coordinates of each district centre where the households live as a proxy for location. A potential concern with attrition is that it can distort sample representativeness and compromise causal inferences when using panel data. However, the ENAHO's overall attrition rate is relatively low, with a re-contact rate of 99.44% among all original households. In addition, the rate of household survey completion is, on average, 89.56% between 2015 and 2019.

3.5.2 Remotely sensed data

The inputs needed to model the river discharge (m^3/s) for basins are watershed, Land Use and Land Cover (LULC), soil data, and Digital Soil Map of the World (DSMW). The details of each geospatial data are explained below.

Watershed

The watershed is a shapefile that delineates the boundaries of the watershed area. In this area, all surface water flows into a common outlet, such as a river, stream, lake, or ocean. Watershed shapefiles typically contain polygons that delineate the boundaries of these drainage basins. The shapefile is obtained from the website GEO GPS PERU[13]. The

---

[12] The ENAHO National Household Survey Dictionary provides detailed explanations of the variables and data collected in the survey. The "ENAHO-2015-2019-100-PANEL" data used in this study contain information from the Housing and Household Characteristics (Module 100) for households included in the 2015-2019 panel. This includes latitude and longitude variables, which were collected in 2018. The document can be downloaded from the following link (see page 24): ENCUESTA NACIONAL DE HOGARES 2018 - Encuesta Continua (datos de panel).

[13] Data can be downloaded from the following link: https://www.geogpsperu.com/2014/02/descargar-cuencas-hidrograficas-del.html.

watershed shapefile is essential for setting up the spatial extent of the SWAT model simulation and for accurately representing the hydrological processes within the watershed.

Digital Elevation Model (DEM)

A DEM is a form of geographic data known as a raster. A raster is a grid of digital, uniform, square cells covering an area on the earth's surface with each cell containing information about elevation (in meters above sea level) and slope (in degrees). I employ ASTER GDEM V3 to model the surface water runoff within the watershed since it offers relatively high-resolution data of land areas on earth of 1 arc second (approximately 30 meters horizontal posting at the equator). These data are downloaded from the National Aeronautics and Space Administration (NASA) [14].

Land Use and Land Cover (LULC)

The LULC is also a raster where each cell corresponds to a specific geographic location and contains information about the land use or land cover type present in that location. Each cell is assigned a discrete value or code representing a particular land use or land cover category[15]. This information is crucial as it provides information about the characteristics of the land surface that influence hydrological processes. For example, impermeable surfaces such as urban areas generate more surface runoff compared to vegetated or forested areas where infiltration rates are higher. By incorporating LULC data into hydrological models, I estimate how much precipitation will directly contribute to runoff, affecting river discharge. The data are downloaded from the Esri website that provides Sentinel–2 10m LULC timeseries from 2017 to 2021[16].

Digital Soil Map of the World (DSMW)

The DSMW is a vector data that provide soil type classification and characterisation of soil parameters. The soil moisture capacity information, that is the soil type and depth, plays a crucial role in regulating runoff generation and river discharge. For instance, wet soils contribute more to runoff generation, whereas dry soils promote infiltration and reduce surface runoff. Hydrological models that integrate DSMW data can simulate soil moisture dynamics over time, helping predict changes in river discharge during wet and dry periods.

[14] Data can be downloaded from the following link: https://search.earthdata.nasa.gov/search.
[15] I use the following class definitions: water, forest, agriculture land, urban area, and barren areas.
[16] Data can be downloaded from the following link: https://www.arcgis.com/apps/instant/media/index.html?appid=fc92d38533d440078f17678ebc20e8e2.

In addition, soil properties influence the partitioning of rainfall into surface runoff and infiltration. Soils with low infiltration rates and high runoff coefficients generate more surface runoff, leading to increased river discharge during storm events. The DSMW data allow modelers to characterise soil properties such as porosity, hydraulic conductivity, and bulk density, which are essential for estimating runoff generation processes and simulating river discharge accurately. The data are downloaded from the Food and Agriculture Organisation of the United Nations (FAO) [17].

Weather data

Weather data are obtained from the NASA – Prediction of Worldwide Energy Resources (NASA-POWER)[18]. This website provides geo-referenced information on global daily terrestrial temperature (maximum and minimum), precipitation (in mm), humidity, wind, and solar radiation at a spatial resolution of 0.5 × 0.5 degrees. Relative humidity, wind speed, solar radiation, and the minimum and maximum air temperatures were obtained from the Climate Forecast System Reanalysis (CFSR), which was designed based on the forecast system of the National Centres for Atmospheric Prediction (NCEP) to provide estimation for a set of climate variability from 1979 to the present day.

VIIRS night-time lights

Floods are idiosyncratic local shocks, occurring generally near a body of water (mostly rivers). This means that flooding affects only small geographic units within a given region, and within the flooded areas, the impact is spatially unequal. Even further, the local population and their assets are also exposed and impacted differently to this natural event.

Night-time lights, as detected by satellites, are increasingly used by economists to proxy for local economic activity and to estimate the impact of weather anomalies on it. For instance, a study undertaken by Felbermayr et al. (2022) uses night-time light emissions to investigate the effects of abnormal weather conditions on economic performance. Other studies, such as Bertinelli and Strobl (2013); Elliott, Strobl, and Sun (2015); and Kocornik-Mina et al. (2020), have studied hurricane, typhoon, and flood impacts on night-time lights imagery, and their

[17] Data can be downloaded from the following link: https://power.larc.nasa.gov/data-access-viewer/.
[18] Data can be downloaded from the following link: https://data.apps.fao.org/map/catalog/srv/eng/catalog.search;jsessionid=1F80A9CA383A4FCDACE7702E1AF73195?node=srv#/metadata/446ed430-8383-11db-b9b2-000d939bc5d8.

findings strengthen the case for assessing the economic impacts of extreme weather at the local level using night-time lights data.

The Defense Meteorological Satellite Program (DMSP) night-time light data have been widely used in the empirical literature as a proxy for predicting GDP, particularly in developing countries, and are valued for their long time-series, spanning from 1992 to 2013. However, these data have several limitations, including image blurring, coarse spatial resolution, lack of calibration, low dynamic range, top-coding, and unrecorded variation in sensor amplification, all of which impair comparability over time and space (Elvidge et al., 2013; Abrahams et al., 2018; Bluhm and Krause, 2018).

In contrast, VIIRS night-time lights data provide greater spatial resolution, better calibration, and improved accuracy in urban areas, capturing finer intra-city variation in light intensity and allowing distinctions between different areas within a city (Gibson et al., 2021). The data are available on a monthly basis, enabling timely, detailed analysis of economic fluctuations and offering accurate proxies for economic density.

In addition, as stated by Gibson et al. (2024), VIIRS night-time lights data offer significant advantages over DMSP for assessing the impact of disasters such as typhoons, because they avoid the excessive spatial autocorrelation seen in DMSP and harmonised datasets and maintain consistent observation times. These features make VIIRS more reliable for detecting localised damage patterns and estimating the true extent and intensity of disaster impacts. Thus, for evaluating the local effects of typhoons and other disasters, VIIRS is a more accurate and robust data source.

As proxy for the local impact of floods, in this study I use the VIIRS night-time lights to capture the economic downturn following the natural event. I extract district level night-time lights data from the VIIRS-Day Night Band (DNB) Cloud Free Monthly Composites (version 1 and tiled) provided by the Earth Observation Group at the Colorado School of Mines[19]. The VIIRS-DNB are spatially and temporally detailed data for night-time lights. In terms of coverage, each VIIRS-DNB image consists of 15-arcseconds grids (463 meters at the equator) and span from 00N latitude to 180W around the entire globe, meaning that the whole of Peru is included.

[19] Data can be downloaded from the following link: https://eogdata.mines.edu/nighttime_light/monthly/v10/.

Following Skoufias et al. 2020, to count for months with no radiance value, I have interpolated between the month before and after (in this case, only three months have no light values).

3.5.3 Government reports: Identifying flooded districts

I use government data to identify the districts in northern Peru affected by floods in 2017. Reports compiled by the National Civil Defence Institute (INDECI)[20] contain information on flood-induced damage at the district level, including the extent of destruction to infrastructure, homes, agricultural land, and the number of affected and displaced individuals. This data enables us to identify the subset of districts impacted by the 2017 floods, forming the treatment group. Among the 57 districts examined, 40 are included in the treatment group, while the remaining 17 constitute the control group.

3.5.4 Construction of the damage index

Empirical economics literature often uses rainfall (e.g., number of days of extreme precipitation) as a source of exogenous variation in explaining the impact of floods on household welfare. More recent studies have used satellite imagery and geocoding, or a combination of both, to create binary measures to identify areas unaffected and affected by flood events that occurred in the past. Less attention has been given to finding a proxy that estimates the (economic) damage caused by flooding due to the data limitations at the local level and its multidimensional nature.

In this paper, the damage associated with flood events is defined as an overall economic impact in the administrative unit (i.e., at the district level) after the occurrence of the disaster. The construction of the damage index involves three stages. First, I use the Soil and Water Assessment Tool (SWAT)[21] to predict daily river discharge within a river basin, using remote sensing data. This provides the volume of water passing specific points along the river. The second and third stages follow the approach of Skoufias et al. (2020), which uses

[20] The Peru's government disaster reports are available freely from the National Institute of Civil Defence (INDECI) of Peru website: https://portal.indeci.gob.pe/informe/informe-de-emergencia/.

[21] Developed by the USDA Agricultural Research Service in collaboration with Texas A&M University and Esri, the SWAT model provides a user-friendly interface for setting up, running, and analysing event-based floods within the ArcGIS software environment. This model has been used to obtain hydrological measures such as evapotranspiration, precipitation, surface runoff, finding that the model captures the daily variability of stream flows across large areas, making it useful for flood simulation.

the time series of streamflow data to calculate the flood intensity[22] in a watershed. Flood intensity is defined as the number of days with high floodwaters exceeding the normal level in the affected area. This intensity measure is then combined with disaggregated local economic activity, obtained from night-time lights data, to estimate an index for flood damage at the district level.

SWAT – input data and model application

The procedures applied from the initial stage to the hydrologic simulation in ArcSWAT – ArcGIS 10.4 are illustrated in Figure 3.A.2 in the Appendix (left panel). The first step of the simulation model involves delineating the basin where the stream or discharge flows and obtain the morphometric parameters[23] of that basin. This is achieved by combining the Watershed shapefile, which contains polygons delineating each river basin in the northern region of Peru, with the Digital Elevation Model (DEM).

Following the delineation step, the process of creating the Hydrologic Response Units (HRUs) begins. The SWAT model allows the delineation of subbasins within a watershed, which are spatially correlated to one another. In each subbasin, I create HRUs which possess unique land use and land cover, and soil attributes. To achieve this, I use an LULC map for the year 2017, derived from Sentinel-2 images Esri, and a global digital map of soil types developed by the Food and Agriculture Organization (FAO). Slope classes ranging from 0% to 20%, 20% to 40%, and steeper than 40% are employed. After defining the input data of this step, it is possible to create the HRUs for each subbasin within a watershed.

The next step involves the creation of tables containing climate variable values such as temperature, precipitation, relative humidity, solar radiation, and wind speed. This is achieved using NASA-POWER data, which provides near real-time weather information at a resolution of 0.5 x 0.5 degrees for the period of 2012 to 2019. Finally, the SWAT model is prepared to simulate variables such as streamflow, evapotranspiration, sediment production and transportation, and nutrient concentration. For this study, daily values of streamflow

[22] For a given level of damage, areas with higher concentration of assets contribute more to the overall losses, as indicated by the level of nocturnal illumination. Consequently, such areas are expected to display significant changes in night-time lights values.

[23] Measurements related to the shape and configuration of river basins, watersheds, and other hydrological features to understand the behaviour of water flow, runoff, and sediment transport in river systems.

(river discharge) are simulated for the period between 2015 and 2019, excluding the initial three years of model warm-up (January 2012–December 2014).

The SWAT simulation of streamflow in ArcGIS can be observed in Figure 3.A.3 in the Appendix. As an example, I present the SWAT simulation for the Tumbes River in ArcGIS (left) and the daily river discharge in m³/s for that river basin (right), which is used to calculate the flood intensity (Equations 2 and 3).

Flood day

Using the daily river discharge (m^3/s) calculated for both each river basin and sub-basin within the period of analysis, I estimate a threshold to define a riverine flood[24] in a day:

$$Flood\ day = 1 \quad if \quad Q > P_{95} + \sigma \tag{2}$$

where P_{95} is the 95th percentile value, and σ is the standard deviation of the river discharge (Q).

Flood intensity

When a flood day is computed, the next step involves calculating the flood intensity. I use the river stream flow as a proxy for intensity, employing the following equation:

$$I_{b,t} = \begin{cases} 0 & : Flood\ day = 0 \\ \dfrac{Q_{b,t} - \bar{Q}_b}{\sigma_b} & : Flood\ day = 1 \end{cases} \tag{3}$$

where $I_{b,t}$ denotes the intensity of a flood event that occurred in the river basin b at date t. $Q_{b,t}$ represents the river discharge in the same basin and time (basin b at date t), while Q_b and σ_b are the mean and standard deviation of river discharge in basin b. The flood intensity is zero if the flood threshold – 95th percentile value plus 1 standard deviation above the average – remains unexceeded[25].

In the context of the flood intensity, normalising the stream flow values within each river basin ensures that the intensity values are on a standardised scale and are independent of the

---

[24] This natural hazard is characterised by the overflow of river water exceeding its normal and expected levels, leading to the inundation of typically dry land in urban and rural areas.

[25] The key assumption consists of two points: i) residents who live close to rivers are prepared for the fluctuations in river water levels, and ii) individuals residing near rivers present a higher level of readiness for variations in river stream flows compared to those living in closer proximity to more stable rivers.

absolute river flows, allowing for more meaningful comparisons and analyses across different basins and time periods. When normalising I constructed two proxies: 1) A simple count variable for the number of flood days within the river basin; and 2) The river discharge in m^3/s registered above the flood threshold[26].

For each month m, I sum the normalised values of flood days to obtain a cumulative value in that month. The cumulative value is a proxy for the total damage inside the river basin.

$$I_{b,m} = \sum_{t=1}^{30} I_{b,t} \ , \quad b = 1, \dots, B; \quad m = 1, \dots, 12 \tag{4}$$

VIIRS night-time lights and flood impact

The monthly average VIIRS night-time lights are used to calculate the weights that aid in finding the locations where flooding had the greatest economic impact and capture the effects of changes in local economic activity. The weights are estimated based on the night-time light cell values within the river basin and those within the district boundaries[27]. They are determined by the following equation:

$$W_{b,d,m,T-1} = \frac{\bar{L}_{b,m,T-1}}{\bar{L}_{d,m,T-1}} \ , \quad T - 1 = 2014, \dots, 2018 \tag{5}$$

where $\bar{L}_{b,m,T-1}$ and $\bar{L}_{j,m,T-1}$ are the mean of VIIRS night-time lights in basin b and in district d, respectively, in month m, and year $T-1$. A one-year lag $(T-1)$, preceding the flood impact, is applied to address concerns regarding endogeneity, specifically reverse causality, between the outcome variables – household income and expenditure per capita – and VIIRS night-time lights. As a robustness test, I also use VIIRS night-time lights data from the year 2013 to validate the main results.

Then, the weight $W_{b,d,m,T-1}$ is multiplied with the flood intensity ($I_{b,m}$) to obtain the monthly flood impact, $FI_{b,d,m}$, in the river basin b, district d, in month m, and year T. The equation is as follows:

$$FI_{b,d,m} = W_{b,d,m,T-1} * I_{b,m} \tag{6}$$

[26] The methodology used to measure the flood intensity is relatively simple and cannot be used to predict local exposure or vulnerability within the river basin.

[27] The study area includes 22 river basins and 57 districts located in the northern region of Peru.

Due to its size, watershed may include more than a single district, and it is assumed that the flood intensity remains consistent across all districts within the river basin.

Flood damage index

As a final step, the flood damage index ($FD_{d,T}$) is calculated by adding all monthly flood impact values ($FI_{b,d,m}$) in river basin b, district d and month m:

$$FD_{d,T} = \sum_{m}^{T} \sum_{b}^{B} FI_{b,d,m}\,, \quad d = 1, \ldots, D; \quad T = 2015, \ldots, 2019 \qquad (7)$$

where T is the sum of all monthly flood impact values in the year and B is the sum of all basins b in district d, and $FI_{b,d,m}$ is the proxy for flood impact from Equation 6. Figure 3.A.2 in the Appendix displays the steps involved in constructing the flood damage index outlined above.

I calculate damage estimates for the period 2017 to 2019 associated with flood events and approximate the economic impact at the district level. In the Appendix, Tables 3.A.3 to 3.A.5 present three groups: 1) the "switcher" districts, 2) the "non-switcher" treated districts, and 3) the "non-switchers" control districts. All groups are ordered from the most to the least damaged districts in 2017.

In addition, Figure 3.2 displays the districts in the northern region of Peru that were most (darker purple) and least affected by floods, based on damage index values. It also shows the sampled households in the panel dataset and the rivers that cross the region.

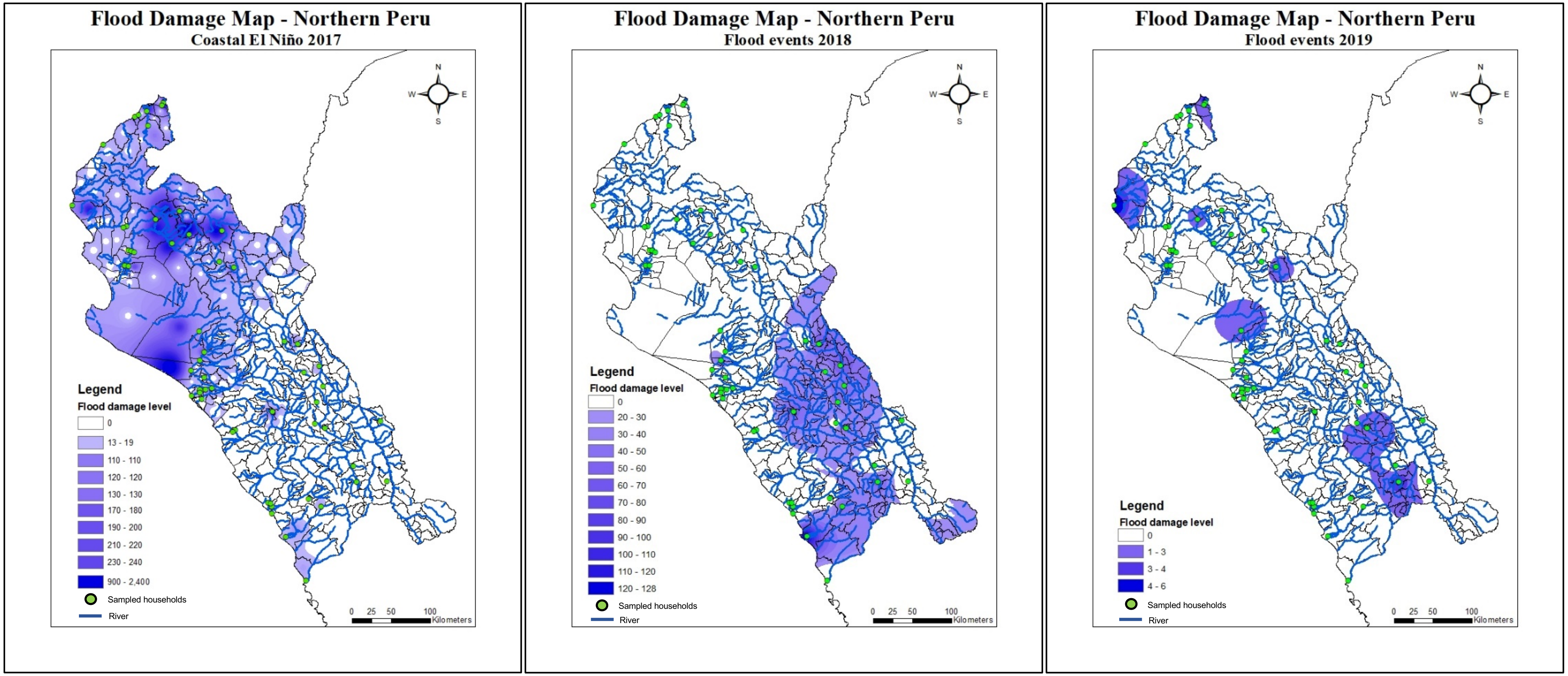


Figure 3.2 Maps of flood damage zones at the district level in northern Peru

Note: The Inverse Distance Weighting (IDW) interpolation method in ArcGIS was used to create the flood maps. These maps illustrate the spatial extent of flood damage hotspots in the northern region of Peru. The continuous smooth surface is classified into categories based on geometric intervals according to the damage level for each year. Flood-affected zones are depicted in deep purple, indicating the highest level of damage, while flood-free areas are shown in white. In addition, green dots denote the sampled households, and the blue line represents the river flowing through the area.
Source: Author’s illustration, 2024.

Descriptive statistics

Table 3.1 provides the summary statistics for all variables used in the analysis below. Note that all continuous variables are measured in logarithmic form, except for age and years of schooling of the household head, and the flood damage index.

A set of main outcome variables described in Panel A includes annual household income and expenditure per capita, and poverty status. I employ fourteen control variables in the regression analysis, which are presented in Panels B and C. For the potential channels analysis, I use income earners, wage income, household savings, and disaster relief in Panels A and B.

Table 3.1 Summary statistics

	All				
	N	Mean	Sd	Min	Max
Panel A: Income, expenditure and related variables					
Annual net household income per capita (in logs)	1,527	8.21	1.04	1.39	12.08
Annual net household expenditure per capita (in logs)	1,530	8.15	0.81	3.40	11.27
Poverty (1,0)	1,530	0.25	0.43	0	1
Urban (1,0)	1,530	0.59	0.49	0	1
Wage income (1,0)	1,293	0.22	0.41	0	1
Household savings (1,0)	1,530	0.44	0.50	0	1
Disaster relief - food and clothing (in logs)	1,530	3.97	2.77	0	9.51
Panel B: Dwelling and household characteristics	1,530	0.57	0.50	0	1
Sewage (1,0)	1,530	0.84	0.36	0	1
Water (1,0)	1,530	0.95	0.21	0	1
Electricity (1,0)	1,525	0.34	0.47	0	1
Concrete roof (1,0)	1,530	0.24	0.43	0	1
Female household head (1,0)	1,530	1.24	0.58	0	2.64
Household size (in logs)	1,530	55.23	15.38	21	97
Age household head (in years)	1,530	7.15	5.05	0	19
Years of schooling household head (in years)	1,526	0.74	0.49	0	2.08
Income earners (in logs)	1,530	0.79	0.41	0	1
Own house (1,0)	1,530	0.57	0.50	0	1
Public health insurance (1,0)					
	1,530	46.25	223	0	2,472.90
Panel C: Damage and geographic characteristics	1,530	0.28	0.45	0	1
Flood damage index	1,530	7.22	1.20	2.87	9.62
Flood dummy (1,0)	1,530	1.76	0.84	0.00	3.48
Distance to near river (in logs)	1,530	4.97	2.05	0.69	8
Terrain slope (in logs)	1,527	8.21	1.04	1.39	12.08
Elevation (in logs)	1,530	8.15	0.81	3.40	11.27

This table presents the summary statistics for the outcome variables and covariates. The household sample is obtained by dropping off-support observations following the application of the PSM approach.
Source: ENAHO, 2015–2019.

Figure 3.3 shows the districts in northern Peru (a total of 34) that switch from the treatment group to the non-treatment group during the period of 2017 to 2019. First, a group of 18 districts moves from the treatment group in 2017 to the non-treatment group in 2018. Then, a group of 16 districts moves from the treatment group in 2018 to the non-treatment group in 2019. Only 6 districts remain in the always-treatment group between 2017 and 2019.

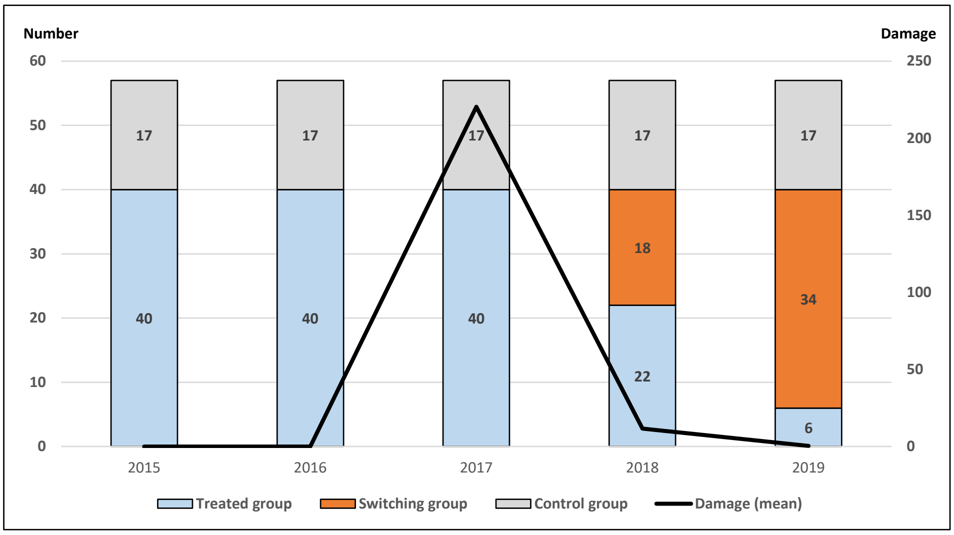


Figure 3.3 Switcher and non-switcher districts from 2015 to 2019

Note: The number of switchers from treated to untreated status was 18 districts in 2018 and 16 districts in 2019. The total number of switchers from 2017 to 2019 was 34 districts.
Source: Author's illustration.

3.6 Pre-estimation

3.6.1 Sample balance analysis

Spatial sorting was not addressed in Chapter 2 because much of the area on which Jakarta is built is prone to flooding, making it difficult to compare observable housing conditions between flood-prone and flood-free areas: (i) approximately 40.95% of DKI Jakarta Province is subject to potential flood inundation (Puspito Sari & Ridhani, 2023); and (ii) during the severe floods of 2002 and 2007, 60% of the city was inundated (World Bank, 2008), reducing the relevance of conducting a sample balance analysis. However, in Chapter 3, this study addresses the potential endogeneity (as a form of measurement error) arising from selectivity or sorting issues. In the context of floods, selection or sorting bias occurs when households may intentionally choose to reside in unaffected districts, thereby influencing the likelihood of experiencing flood impacts. For instance, in response to this natural event, a household might choose to relocate or rebuild in low- or non-risk areas to mitigate or avoid the negative effects on their income, consumption, wealth and work. Therefore, estimates relative to the impact of floods on the household welfare could be understated. If this issue is not appropriately addressed, it can lead to biased estimates.

To identify any potential selection bias in the dataset, I perform two tests. First, I evaluate whether areas unaffected and affected by floods differed in terms of socioeconomic characteristics ex-ante. Specifically, the test indicates whether the two groups differ systematically on observable characteristics. Second, I use a multiple linear regression (OLS) of the treatment dummy on pre-determined covariates to detect the possibility of systematic sorting.

Table 3.2 Balance tests between unaffected and affected areas – before PSM

	Control Group	Treatment Group	Single difference	N	p-value	
Years: 2015/2016 Dependent variable	(1)	(2)	(3)	(4)	(5)	
Sewage (1,0)	0.564	0.542	0.022	629	0.637	
Water (1,0)	0.826	0.806	0.019	629	0.601	
Electricity (1,0)	0.987	0.906	0.080	629	0.001	***
Concrete roof (1,0)	0.383	0.326	0.056	627	0.207	
Female household head (1,0)	0.221	0.258	-0.037	629	0.365	
Household size (in logs)	1.272	1.232	0.041	629	0.459	
Age of household head (in years)	54.409	53.788	0.622	629	0.670	
Years of schooling of household head (in years)	6.383	7.254	-0.872	629	0.067	*
Income earners (in logs)	0.734	0.706	0.028	629	0.545	
Ownership of house (1,0)	0.772	0.783	-0.012	629	0.767	
Public health insurance (1,0)	0.664	0.531	0.133	629	0.004	***

Note: Columns 1 and 2 report the sample mean for the unaffected (control) and affected (treatment) groups, respectively. The data is restricted to 2015 and 2016 household survey. Column 3 shows single difference in means between non-affected and districts affected by floods. Column 4 displays the number of observations, while column 5 presents the p-value on each balance test. ∗∗∗p < 0.01, ∗∗ p < 0.05, ∗ p < 0.1.
Source: ENAHO, 2015–2019.

In Table 3.2, column 3 shows single differences in means of key variables between affected and unaffected areas using the first two years of the panel data (2015 and 2016). Three of these differences are statistically significant at 1% and 10% (electricity, years of schooling of household head, and public health insurance), indicating unbalanced household characteristics between the two groups. In addition, in Table 3.3, column 6, the OLS regression presents one variable is associated with the treatment dummy with a 5% significance level. The difference in statistical significance for years of schooling between the two tables stems from how standard errors are calculated under each method. Despite this, I conclude there is a sorting issue in the dataset.

Table 3.3 Exogeneity tests: OLS regressions – before PSM

	OLS	N	p-value	
Years: 2015/2016 Dependent variable	(1)	(2)	(3)	
Sewage (1,0)	-0.022	629	0.865	
Water (1,0)	-0.019	629	0.889	
Electricity (1,0)	-0.080	629	0.040	**
Concrete roof (1,0)	-0.056	627	0.519	
Female household head (1,0)	0.037	629	0.567	
Household size (in logs)	-0.041	629	0.688	
Age of household head (in years)	-0.622	629	0.757	
Years of schooling HH head (in years)	0.872	629	0.330	
Income earners (in logs)	-0.028	629	0.713	
Ownership of house (1,0)	0.012	629	0.874	
Public health insurance (1,0)	-0.133	629	0.139	

Note: This table displays the correlation between the dummy variable for affected and unaffected areas and a set of covariates. Each row represents a separate OLS regression of the dependent variable (covariate) on the dummy variable (equal to 1 if treated, 0 otherwise). All OLS regressions include district fixed effects. Columns 1 and 2 display the coefficient related to the covariate and the number of observations, while column 3 displays the p-value. Data are restricted to 2015 and 2016 household survey. $***p < 0.01$, $** p < 0.05$, $* p < 0.1$.
Source: ENAHO, 2015–2019.

By applying propensity score matching to the dataset, I find observations without a counterfactual pair (off-support) as shown in Figure 3.4. These observations are dropped, and after several iterations, the matching method of kernel with uniform distribution is chosen to obtain a balanced dataset.

To address the endogeneity, I employ propensity score matching during the pre-estimation step. This technique enables us to create counterfactuals for households living in flooded districts within the control group of the dataset. The aim is to achieve a balanced condition where the characteristics of observations in flooded areas are similar to those in non-flooded areas.

I also display kernel density plots to see how covariates overlap before and after matching. In Figure 3.5, there is strong overlap in the distribution of propensity scores between the treated and untreated groups.

As a result, matching based on propensity scores effectively balance the two groups on observed covariates, leading to unbiased estimates of the treatment effect.

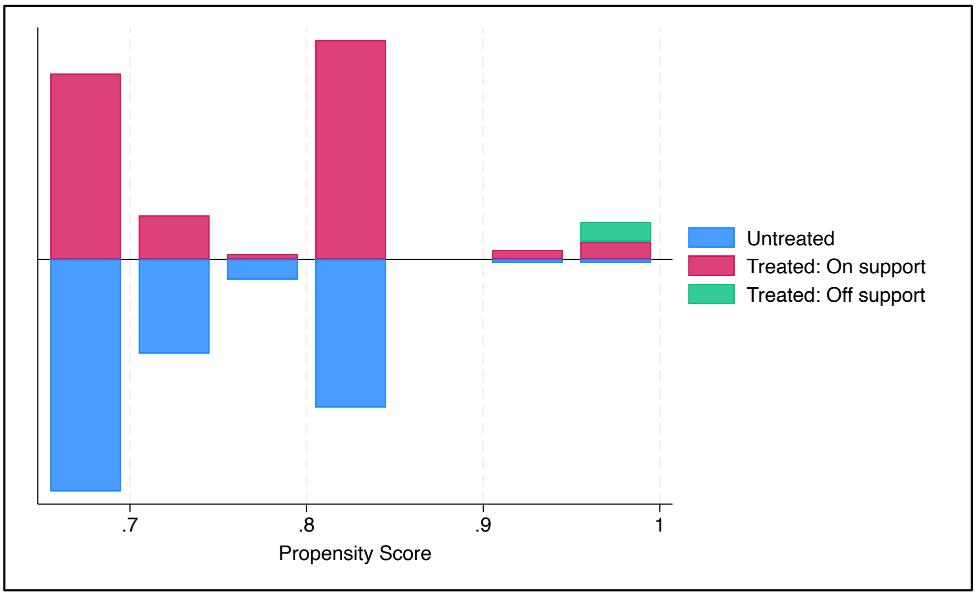


Figure 3.4 Propensity score matching

Note: The figure displays the distribution of propensity scores for the treated and control groups after matching, using the kernel matching method with a uniform distribution. Blue and red bars represent the matched pairs, where covariates are balanced across treated and control groups. Observations without a counterfactual pair (off-support), which are excluded from the analysis, are shown in green.
Source: ENAHO, 2015–2019.

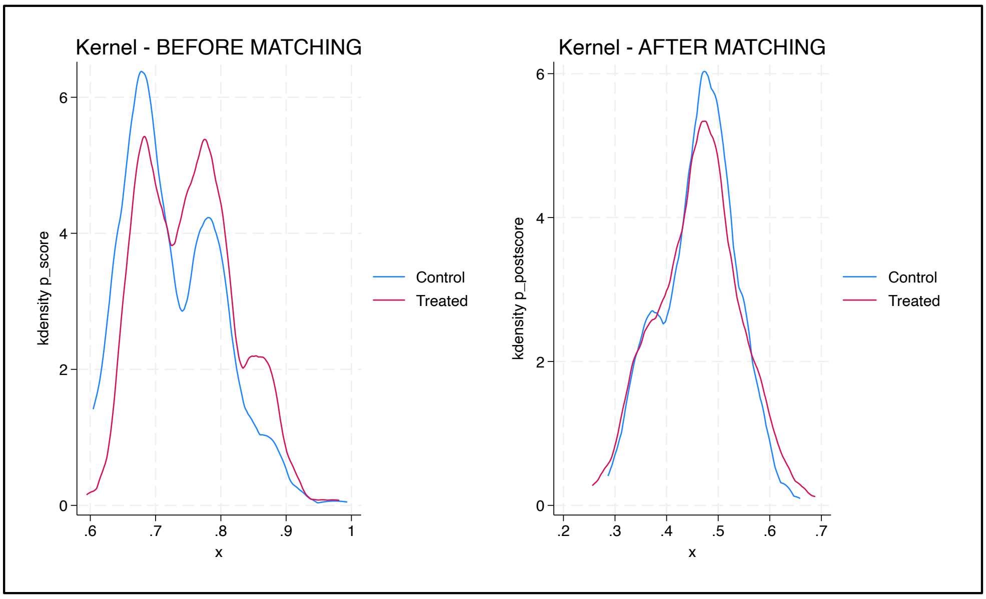


Figure 3.5 Kernel density before and after matching

Note: The figure illustrates the improved balance of propensity scores between the treated and control groups after matching. The kernel density distribution for the control group is shown in blue, while that of the treated group is in red.
Source: ENAHO, 2015–2019.

Next, I re-evaluate the association between the treatment dummy and these four variables using matched data and incorporating weights obtained from propensity score matching[28]. The new weights used in this estimation are calculated by multiplying the sample weights and the propensity score matching weights. The results show that the associations of the three variables (electricity, years of schooling household head, and public health insurance) are no longer significant; however, the household size variable still has a significant association albeit weaker significance (see Table 3.4, column 6). A joint significance test further shows that the observed covariates do not explain the treatment dummy.

[28] The Propensity Score Matching technique calculates weights (propensity scores) for each observation in the control group to match with observations in the treatment group, based on their observed characteristics (covariates). In this study, the weights assigned to observations in the control group present different values within the pre-treatment period (2015 and 2016). Then, I use the weights for the control group calculated in 2016 and duplicate them for the years 2017, 2018, and 2019. I follow this procedure since the event study assumes parallel trends in both the pre- and post-treatment periods.

Table 3.4 Balance and exogeneity tests between unaffected and affected groups – after PSM

	Control Group	Treatment Group	Single difference	p-value		OLS	p-value	N
Years: 2015/2016 Dependent variable	(1)	(2)	(3)	(4)		(5)	(6)	(7)
Sewage (1,0)	0.717	0.688	0.029	0.623		-0.029	0.787	610
Water (1,0)	0.870	0.858	0.012	0.800		-0.012	0.842	610
Electricity (1,0)	0.953	0.949	0.004	0.917		-0.004	0.926	610
Concrete roof (1,0)	0.416	0.371	0.045	0.546		-0.045	0.618	608
Female household head (1,0)	0.198	0.280	-0.083	0.199		0.083	0.317	610
Household size (in logs)	1.400	1.234	0.166	0.055	*	-0.166	0.126	610
Age of household head (in years)	54.235	55.755	-1.520	0.480		1.520	0.467	610
Years of schooling household head (in years)	7.193	7.715	-0.522	0.562		0.522	0.666	610
Income earners (in logs)	0.839	0.747	0.092	0.239		-0.092	0.271	610
Ownership of house (1,0)	0.756	0.807	-0.051	0.361		0.051	0.574	610
Public health insurance (1,0)	0.447	0.509	-0.062	0.384		0.062	0.523	610
Test of joint significance			F-stat: 1.53 (p-value: 0.1464)					

Note: Columns 1 and 2 report the weighted mean for the control and treatment groups, respectively. Column 3 shows the weighted single difference in means between districts unaffected and affected by floods. Column 5 displays the estimates of the treatment dummy on household characteristics using an OLS regression with weights. A test for the joint significance of the coefficients is performed after running an OLS regression of the treatment dummy on the base line covariates. Data are restricted to 2015 and 2016 household survey. ***p < 0.01, **p < 0.05, *p < 0.1.
Source: ENAHO, 2015–2019.

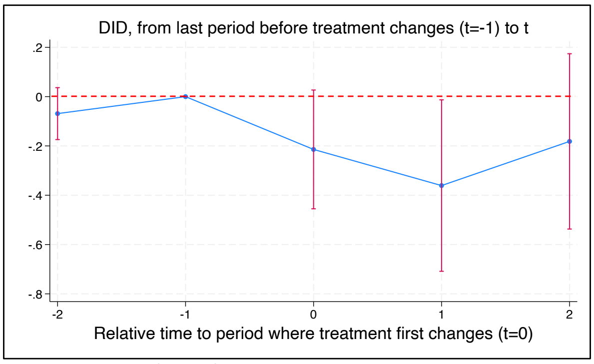
The matching process results in the discarding of several observations, leaving a final sample of 1530 observations (from a total of 1614 observations). I now expand the dataset based on the new weights constructed by sample weights and matching results, as most of the covariates are no longer significantly associated with the treatment variable (balanced conditions achieved). As a result, this dataset has an unbalanced panel data structure.

3.6.2. Event study assumptions analysis

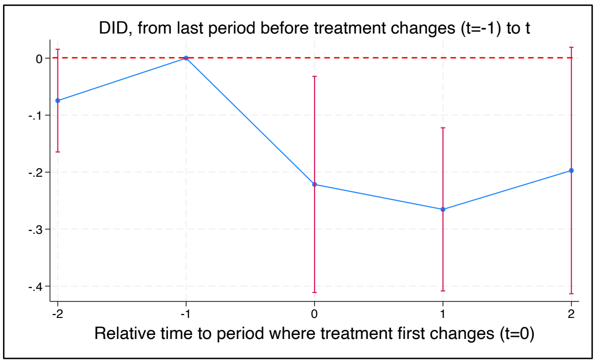
In this section, I test the parallel trends and no anticipation assumptions using the balanced panel data obtained from the pre-estimation analysis. On one hand, statistical tests for pre-treatment trends cannot be employed in this study, as the matched data only includes one year (i.e., 2015) prior to the treatment year (i.e., 2017). However, it is assumed that if the differences in pre-trends between the treated and control groups (also referred to as placebo effects) are not statistically significant in the regression, and if the confidence interval for the pre-treatment mean effect crosses zero in the event study plot, I can expect that the trends would have continued in the treated group had it not received the treatment. On the other hand, the difference in outcomes between the treated and control groups is not significantly different before the treatment year, suggesting that there were no pre-existing household responses indicating anticipation of the treatment.

Parallel trends

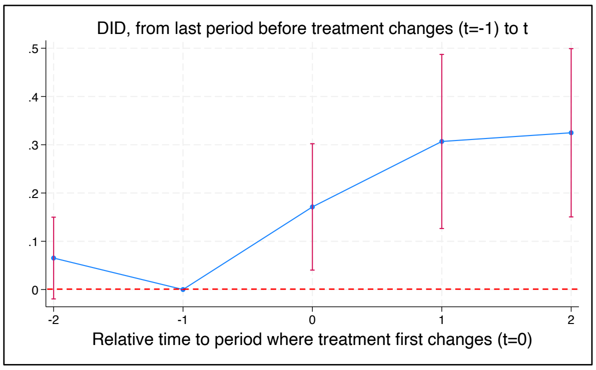
The diff-in-diff design in this study involves two pre-treatment years (2015–2016) and another three years post-treatment (2017–2019). As the floods occurred in 2017, year 2016 is dropped and serves as the base period against which all effects are compared. This reduces to only one lead with respect to the pre-treatment period (i.e., year 2015). I include a set of covariates in the model because they are important and strongly affect the outcomes to help ensure the parallel trends assumption holds. As stated by Caetano and Callaway (2024), adjusting for observable characteristics increases the likelihood that the parallel trends assumption holds, leading to more accurate and reliable ATT estimates. This requires that the covariates strongly influence the outcomes and evolve in a non-parallel fashion across treatment and control groups. The covariates included in the current analysis have been used in previous studies (i.e., dwelling and demographic characteristics), and evolve differentially across treatment and control groups over time, as can been seen in Figure 3.A.4.



Income per capita (in logs)



Expenditure per capita (in logs)



Poverty (1,0)

Figure 3.6 Event study plots for impact of floods on household welfare and poverty

Note: Estimation results based on de Chaisemartin et al. (2023) estimator, clustered standard errors, 95% CI. Negative numbers indicate the years before the coastal El Niño in 2017, and non-negative numbers represent years after the coastal El Niño in 2017.
Source: ENAHO, 2015–2019.

Figure 3.6 illustrates the parallel trends in the pre-treatment (Effect –2, year 2015). The coefficients are not statistically significant for household income and expenditure per capita, and poverty status, indicating that the assumption holds.

No anticipation

The presence of no anticipation effects implies that households located in flooded districts do not alter their behaviour in anticipation of the impact of floods. When anticipation effects are present, households may adjust their behaviour before the floods occur, expecting their effects to manifest. To test for anticipation effects, I examine if the pre-existing differences prior to the occurrence of the flood events (years 2015 and 2016) changed over time. The double differences in Table 3.5 are not statistically significant, implying that households were unlikely to anticipate and act on the shock ex-ante, at least at the district level.

Table 3.5 Double differences between treated and control groups

Years: 2015/2016 Dependent variable	Control Group (1)	Treatment Group (2)	Double difference (3)	N (4)	p-value (5)
Sewage (1,0)	0.717	0.688	-0.013	610	0.785
Water (1,0)	0.870	0.858	-0.027	610	0.650
Electricity (1,0)	0.953	0.949	0.040	610	0.563
Concrete roof (1,0)	0.416	0.371	-0.011	608	0.865
Female household head (1,0)	0.198	0.280	0.032	610	0.336
Household size (in logs)	1.400	1.234	-0.071	610	0.154
Age of household head (in years)	54.235	55.755	1.358	610	0.277
Years of schooling HH head (in years)	7.193	7.715	-0.010	610	0.972
Income earners (in logs)	0.839	0.747	0.002	610	0.985
Ownership of house (1,0)	0.756	0.807	0.171	610	0.107
Public health insurance (1,0)	0.447	0.509	-0.008	610	0.930

Note: Column 1 and column 2 report the sample mean for the control and treatment groups, respectively. The data are restricted to the 2015 and 2016 household surveys. Column 3 displays the estimate of the interaction term between the treatment dummy and the year 2016, using a Difference-in-Difference regression of each variable on interaction with weights. ***p < 0.01, **p < 0.05, *p < 0.1.
Source: ENAHO, 2015–2019.

Another indicator supporting the assumption of no anticipation is that residents failed to relocate to safer areas after the flood events in 2017. Table 3.6 shows that the number of household members born in the same district remained unchanged after the flood occurrence. Moreover, the percentage of households owning a second house remained low, accounting for 7.1% of the total surveyed over the analysis period.

Table 3.6 District of birth and households owing a second house – number of households

District of birth	2015	2016	2017	2018	2019
Same	177	182	185	183	185
Different	125	126	122	124	121
Number of observations	302	308	307	307	306

Household owns 2nd house	2015	2016	2017	2018	2019
Yes	17	28	27	15	24
No	285	280	280	292	282
Number of observations	302	308	307	307	306

Source: ENAHO, 2015–2019.

The above results, therefore, produce supporting evidence of the causal impact of flood shocks on households in the districts affected.

3.7 Results

3.7.1 Main results

I begin by estimating equation (1) for net annual household income and expenditure, measured in log per capita terms, and poverty status as a binary variable. The ENAHO determines whether a household is poor based on a comparison of their income or expenditure per capita to the national poverty line. Table 3.7 presents the results of the estimations. When analysing the local economic impact of floods, I find that they reduced household income and expenditure per capita and increased the poverty level among the households located in northern Peru.

In column (1) of Table 3.7, the coefficient on the binary switching treatment is estimated to be –0.393. This implies that the treated group's income decreased by 32.5% relative to the average income of non-flooded households in the same region. Similarly, in column (2), the coefficient is –0.356, which translates to a 29.9% average decline in expenditure among the treated group relative to those outside the flooded areas. These findings are also confirmed by using the continuous switching treatment – the coefficients are also negative (higher in magnitude than using continuous variable) and translate to a 15.9% and 14.5% decline in income and expenditure per capita, respectively, relative to the unflooded control group[29].

[29] These percentage changes are calculated based on the average of the flood damage index, which is 46.47. The %change = exp ($\boldsymbol{\tau_e}$*46.47) − 1.

In monetary terms, the loss of income and expenditure per capita is PEN 1,514.31 (USD 464.41) and PEN 1,154.12 (USD 353.94), respectively, following the flood events.

On the other hand, the poor are usually the most vulnerable to the effects of floods (Richie et al., 2020), resulting in a decrease in income and assets and further entrenching their poverty (Carter et al., 2007; World Bank, 2013; Hallegatte et al., 2016). As presented in column (3) of Table 3.9, I find a similar pattern in the northern region of Peru, where households in flooded areas were 41.74 percentage points more likely to fall below the poverty line relative to those outside the affected areas. Using the flood damage index, poverty levels rose by 19 percentage points in areas impacted by floods[30].

Table 3.7 Impact of flood shocks on household welfare and poverty

Dependent variable:	**Income per capita (in logs)**		**Expenditure per capita (in logs)**		**Poverty (1,0)**	
A: Binary treatment						
Effect -2 (year 2015)	-0.069		-0.075		0.065	
	(0.054)		(0.046)		(0.043)	
Effect 0 (year 2017)	-0.214	*	-0.222	**	0.171	**
	(0.123)		(0.097)		(0.067)	
Effect 1 (year 2018)	-0.361	**	-0.265	***	0.307	***
	(0.177)		(0.073)		(0.092)	
Effect 2 (year 2019)	-0.182		-0.197	*	0.325	***
	(0.181)		(0.110)		(0.089)	
ATT (flooded district=1)	**-0.393**	*	**-0.356**	***	**0.417**	***
	(0.220)		**(0.111)**		**(0.111)**	
B: Flood damage level						
ATT (flooded district=index)	**-0.004**	*	**-0.003**	***	**0.004**	***
	(0.002)		**(0.001)**		**(0.001)**	
District FE	Yes		Yes		Yes	
Year FE	Yes		Yes		Yes	
Cluster (district)	Yes		Yes		Yes	
Controls	Yes		Yes		Yes	
N	913		913		913	

The dependent variables are listed across the top row. Income and expenditure per capita are measured as the log of net annual household income and expenditure per member, respectively. Poverty is a binary variable that takes on a value of one if the household is below the poverty line, else zero. Estimation results are based on the de Chaisemartin et al. (2023) estimator, controlling for dwelling, household and geographic characteristics, as well as district and year fixed effects. Standard errors are clustered at the district level.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.
Source: ENAHO, 2015–2019.

[30] Using the average damage index of 46.47, the expected impact on the probability of falling below the poverty line is calculated as: $Pr = (\tau_e * 46.47) * 100$.

In addition, there are dynamic treatment effects in the aftermath of the coastal El Niño floods. Significant economic impacts on income and expenditure per capita can be observed between 2017 and 2019 (mostly at the 5% significance level), but they dissipate for income per capita in the following year. In the case of poor households, the treatment effects persist for up to three years (at the 1% and 5% significance level).

3.7.2 Robustness test

In this section, I apply the Dynamic TWFE regression model using equation (1). Unlike the estimator proposed by de Chaisemartin et al. (2023), which can model a treatment variable that switches on and off, this technique assumes that once the treatment (flood shocks) begins for a group (affected households), it continues throughout the remaining time periods (2018–2019). Although this assumption simplifies the robustness test of the main results in this paper, it may raise problems related to negative weighting (Roth and Sant'Anna, 2023). Therefore, the estimates from this specification should be interpreted cautiously.

Alternative methods for estimating treatment effects under quasi-random treatment have been developed by Roth and Sant'Anna (2023) and Callaway and Sant'Anna (2021). These estimators are designed for settings with staggered treatment and heterogeneous effects across groups and time, and they require strictly balanced panel data. Another relevant method is proposed by Sun and Abraham (2021), who also adopt a staggered difference-in-differences model. Unfortunately, all three methods are infeasible in this study. The group of districts affected by the major flood shock in 2017 includes both switchers and always–treated in the subsequent years (2018 and 2019), and the panel data are unbalanced.

As such, I employ the Dynamic TWFE estimator despite the risk of negative weighting, which can yield biased estimates even when the treatment variable is plausibly exogenous.

Table 3.8 presents the Dynamic TWFE regression using a binary treatment (flooded district = 1). While the signs of the coefficients are consistent with those reported in the previous results section (Table 3.7), the magnitudes and significance levels vary. In particular, the dynamic TWFE specification produces more precise estimates and higher statistical significance for several outcomes, notably poverty. These results reinforce the conclusion that floods had negative effects on household welfare and increased the incidence of poverty in affected districts.

Table 3.8 Impact of flood shocks on household welfare and poverty, Dynamic (TWFE)

Dependent variable:	**Income per capita (in logs)**		**Expenditure per capita (in logs)**		**Poverty (1,0)**	
Binary treatment						
Effect -2 (year 2015)	-0.066		-0.051		0.145	***
	(0.099)		(0.073)		(0.051)	
Effect 0 (year 2017)	-0.143		-0.227	***	0.225	***
	(0.096)		(0.072)		(0.053)	
Effect 1 (year 2018)	-0.294	**	-0.273	***	0.329	***
	(0.128)		(0.072)		(0.084)	
Effect 2 (year 2019)	-0.056		-0.165	*	0.294	***
	(0.133)		(0.090)		(0.064)	
ATE (flooded district=1)	**-0.165**	*	**-0.222**	***	**0.283**	***
	(0.090)		**(0.055)**		**(0.056)**	
District FE	Yes		Yes		Yes	
Year FE	Yes		Yes		Yes	
Cluster (district)	Yes		Yes		Yes	
Controls	Yes		Yes		Yes	
N	1,521		1,521		1,521	

The dependent variables are listed across the top row. Income and expenditure per capita are measured as the log of net annual household income and expenditure per member, respectively. Poverty is a binary variable that takes on a value of one if the household is below the poverty line, else zero. Estimation results are based on the Dynamic TWFE estimator, controlling for dwelling, household and geographic characteristics, as well as district and year fixed effects. Standard errors are clustered at the district level.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.
Source: ENAHO, 2015–2019.

The average effect of floods on net household income and expenditure per capita is –0.165 and –0.222, which translate to a 15.2% and 19.9% decline, respectively. For poverty, the average effect is 0.283, indicating a 28.3 percentage points increase in its incidence among households.

In addition, Figure 3.7 plots the coefficients from the separate regression of household income and expenditure per capita. The coefficients prior to 2016 are statistically insignificant (except for poor households), implying that the parallel trends assumption holds. Graphically, the results using the heterogeneity robust estimator are basically identical and follow the same dynamic treatment effects as the main results using the traditional TWFE specification.

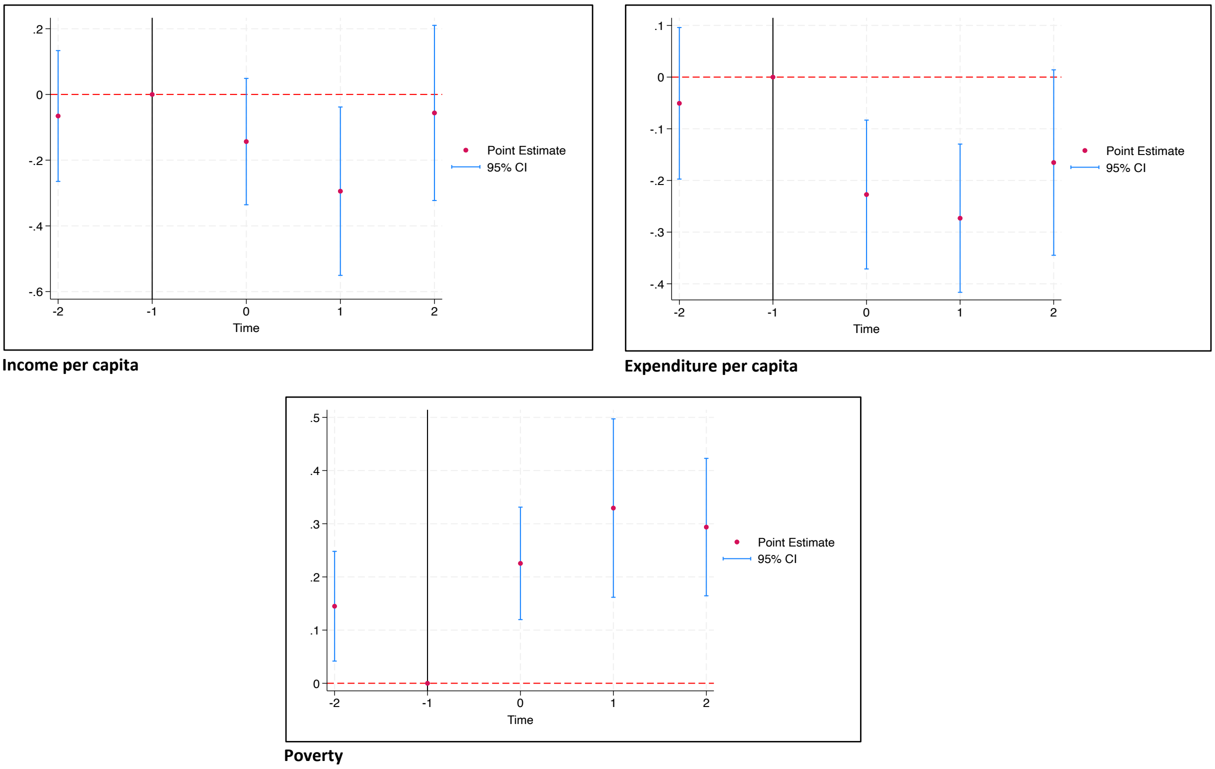


Figure 3.7 Event study plot for impact of floods on household welfare and poverty

Note: Estimation results based on Dynamic TWFE, clustered standard errors, 95% CI. Negative numbers indicate the years before the coastal El Niño in 2017, and non-negative numbers represent years after the coastal El Niño in 2017.
Source: ENAHO, 2015–2019.

Results using weights based on 2013 VIIRS night-time lights

As noted above, the flood damage index is employed to validate the findings under the binary switching treatment. However, concerns about reverse causality might arise, as VIIRS night-time lights weights are included as an input into the explanatory variable of flood intensity, even though they are also an outcome that could be correlated with the main dependent variables such as household welfare and poverty status. To examine the extent to which the empirical results may be affected by this choice, I construct an alternative flood damage index using weights based on 2013 VIIRS night-time lights and re-estimate equation (1). Note that the construction of the index follows the same procedure outlined in Section 3.5.4, but exclusively uses VIIRS night-time lights data from 2013 instead of the 2014–2018 data.

Table 3.A.6 in the Appendix presents the results, which are similar in both sign and significance levels but show larger magnitudes compared to those in Table 3.9. Specifically, income and expenditure per capita declined by 28.5% (against 15.9%) and 26.2% (14.5%), respectively, while households in affected areas experienced an increase in poverty status of 36.8% (19%).

3.7.3 Extended analysis

In this section, I expand the empirical investigation by examining the heterogeneity of the main results related to household income and expenditure per capita to the extent of urban and rural contexts. I conduct this analysis using sub-samples of data that cover populations in both areas. In columns (1) and (2) of Table 3.9, floods did not affect household income in urban and rural areas. The absence of a significant effect may be attributed to the region's commencement of recovery from the flood effects through job creation in affected areas. On the other hand, only households located in urban areas experienced a decrease in expenditure due to the flood disasters (29.2%). This may be explained by an increase in spending on housing repair and furniture for flooded households. In general, poverty increases in the affected areas, however, it is concentrated more in urban areas compared to rural areas, as shown in column (3) of Table 3.9. This may be explained by the slow process of economic recovery after the flood event, especially in promoting job recovery, which can be challenging due to damaged infrastructure, displaced people, and heavy casualties and property loss.

Using household net income, I create quartiles to assist in determining whether I observe heterogenous impacts of flooding across different socioeconomic status. I divide households using their corresponding income into quartiles representing poor (Q1), and rich (Q4) households. Then, I estimate the causal impact of floods on income per capita. provides a summary of results for the coefficient of interest separately for each group. I report only the coefficient on the flood impact. Results are presented in Columns (5) and (6) of Table 3.9. Although I see no significant difference in incomes across the two extreme quartiles, the results are still instructive. I observe that poor households experience a decline in their income per capita, while richer households face an increase in their income.

Table 3.9 Heterogeneity impacts

Dependent variable	(1) Income per capita (in logs)		(2) Expenditure per capita (in logs)		(3) Poverty (1,0)		(4) Socioeconomic Status (S/)	
	Urban	**Rural**	**Urban**	**Rural**	**Urban**	**Rural**	**Quartile 1 Poorest**	**Quartile 4 Richest**
A: Binary treatment								
Effect -1 (year 2015)	-0.026	0.147	-0.017	-0.078	0.051	0.133	-0.307	-0.209
	(0.059)	(0.098)	(0.049)	(0.074)	(0.034)	(0.095)	(0.222)	(0.131)
Effect 0 (year 2017)	-0.010	-0.341	-0.212 **	-0.100	0.116	0.218	-0.419	0.207
	(0.160)	(4.040)	(0.096)	(3.467)	(0.093)	(4.182)	(0.598)	(0.337)
Effect 1 (year 2018)	-0.213	-0.509	-0.271 ***	-0.191	0.315 ***	0.179	-0.038	-0.057
	(0.254)	(4.071)	(0.087)	(3.453)	(0.120)	(4.226)	(0.762)	(0.374)
Effect 2 (year 2019)	0.015	-0.245	-0.168	-0.254	0.311 ***	0.308	-0.926	-0.079
	(0.253)	(2.142)	(0.143)	(2.006)	(0.111)	(2.537)	(0.924)	(0.433)
ATT (flooded district=1)	**-0.110**	**-0.541**	**-0.345** **	**-0.268**	**0.394** ***	**0.348**	**-0.580**	**0.033**
	(0.324)	**(4.973)**	**(0.137)**	**(4.233)**	**(0.147)**	**(5.173)**	**(0.929)**	**(0.547)**
B: Flood damage level								
ATT (flooded district=index)	**-0.001**	**-0.003**	**-0.005** **	**-0.001**	**0.005** ***	**0.002**	**-0.003**	**0.000**
	(0.004)	**(0.027)**	**(0.002)**	**(0.023)**	**(0.002)**	**(0.029)**	**(0.004)**	**(0.010)**
N	539	374	539	374	539	374	177	216
District FE	Yes		Yes		Yes		Yes	
Year FE	Yes		Yes		Yes		Yes	
Cluster (district)	Yes		Yes		Yes		Yes	
Controls	Yes		Yes		Yes		Yes	

Note: The dependent variables are listed in the top row. Income and expenditure per capita are dummy variables equal to one for urban households and zero for rural households. Poverty is a dummy variable equal to one for households below the poverty line in urban areas and zero for households below the poverty line in rural areas. Socioeconomic status ranges from one (poorest) to four (richest). Estimation results are based on the de Chaisemartin et al. (2023) estimator, controlling for dwelling, household and geographic characteristics, as well as district and year fixed effects. Standard errors are clustered at the district level. ∗∗∗ $p < 0.01$, ∗∗ $p < 0.05$, ∗ $p < 0.1$. Source: ENAHO, 2015–2019.

In addition, Figure 3.8 displays the dynamic treatment effects in urban areas after the occurrence of flood events. There are short- and long-run causal effects on poverty that remain higher compared to the years prior to the floods occurring.

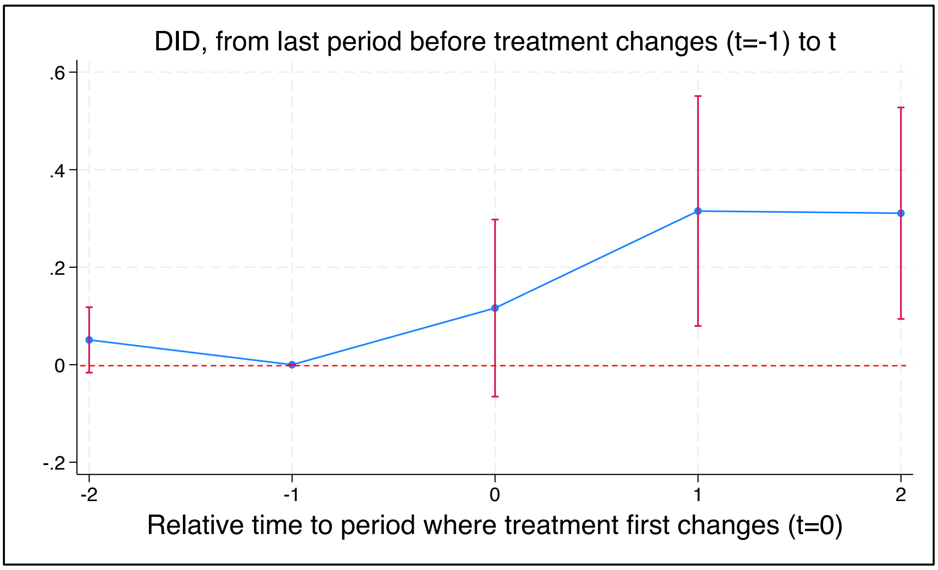


Figure 3.8 Economic dynamics after the disaster on poverty in urban areas

Note: Estimation results based on de Chaisemartin et al. (2023) estimator, clustered standard errors, 95% CI. Negative numbers indicate the years before the coastal El Niño in 2017, and non-negative numbers represent years after the coastal El Niño in 2017.
Source: ENAHO, 2015–2019.

3.7.4 Potential channels

Previous estimation results show a decline in income and expenditure of the treatment group relative to the control due to flooding. Here, I consider the potential channels behind these observed impacts. I first characterise the potential pathways with reference to the above empirical findings and test each possible path using the event study regression analysis.

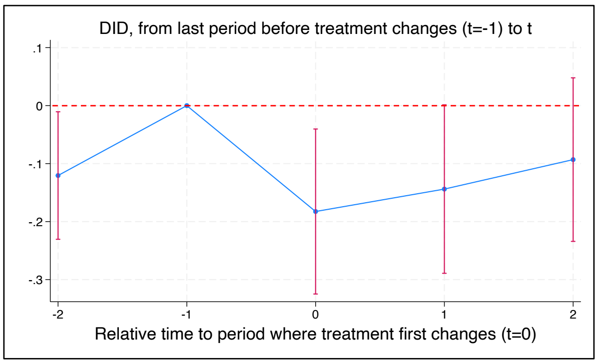
Flooding may drive a decline in income by (i) reducing the number of income earners within the household, and (i) deteriorating wage income. Nevertheless, households could engage in consumption smoothing through (iii) depleting household savings, and (iv) donations or disaster relief to support their recovery from disasters. Table 3.10 displays that floods have a negative impact on household income earners, wage income and savings, while showing positive effects on disaster relief – such as food and clothing – provided to households.

Table 3.10 Income earners, wage income and disaster relief

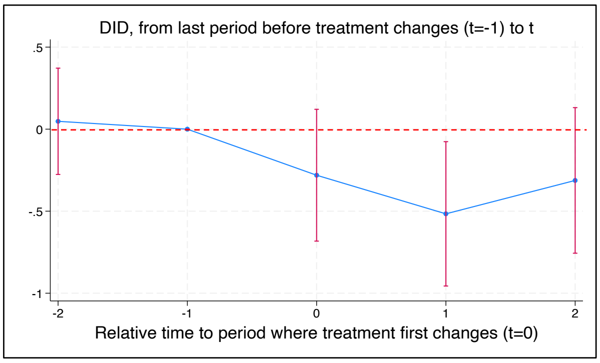
Dependent variable:	Number income earners (in logs)		Wage income (1,0)		Household savings (1,0)		Disaster relief - food and clothing (in logs)	
A: Binary treatment								
Effect -2 (year 2015)	-0.121	**	0.048		-0.145	***	0.446	
	(0.056)		(0.165)		(0.045)		(0.363)	
Effect 0 (year 2017)	-0.183	**	-0.281		0.006		1.214	**
	(0.073)		(0.205)		(0.130)		(0.509)	
Effect 1 (year 2018)	-0.144	*	-0.516	**	-0.255	*	0.939	
	(0.074)		(0.224)		(0.145)		(0.620)	
Effect 2 (year 2019)	-0.093		-0.312		-0.425	***	0.885	
	(0.072)		(0.226)		(0.132)		(0.726)	
ATT (flooded district=1)	**-0.218**	**	**-0.602**	**	**-0.351**	*	**1.579**	**
	(0.098)		**(0.269)**		**(0.191)**		**(0.789)**	
B: Flood damage level								
ATT (flooded district=index)	**-0.002**	**	**-0.005**	**	**-0.003**	***	**0.015**	**
	(0.001)		**(0.002)**		**(0.002)**		**(0.006)**	
District FE	Yes		Yes		Yes		Yes	
Year FE	Yes		Yes		Yes		Yes	
Cluster (district)	Yes		Yes		Yes		Yes	
Controls	Yes		Yes		Yes		Yes	
N	913		436		913		171	

The dependent variables are listed across the top row. The number of income earners is measured as the log of household members who earn income. Wage income is a dummy that takes on a value of one if the household head earns a wage as an independent contractor, else zero. Household savings is a dummy that takes on a value of one if the household head has a bank account, else zero. Disaster relief is measured as the log of the monetary value of food and clothing donations. Estimation results are based on the de Chaisemartin et al. (2023) estimator, controlling for dwelling, household, and geographic characteristics, as well as district and year fixed effects. Standard errors are clustered at the district level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.
Source: ENAHO, 2015–2019.

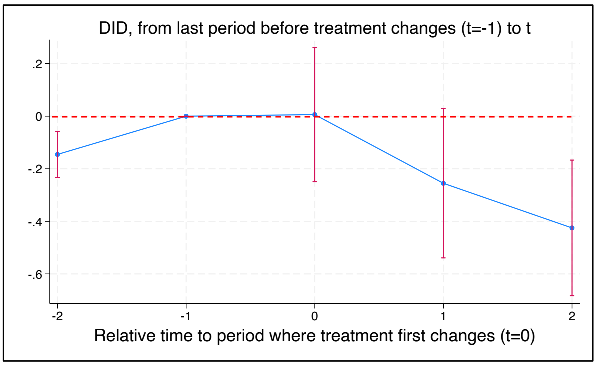
I find that household workers are more likely to lose their jobs following a natural shock. This could be explained by the disruption of local economic activity. For instance, if one family member experiences income loss, their revenue may decline by 19.6% due to the flood. Additionally, households that experienced larger and longer floods during the period 2017 to 2019 may have seen much larger decreases in wage income and household savings compared to floods of shorter duration. Figure 3.9 illustrates the dynamic treatment effects of floods on both income earners and wage income.



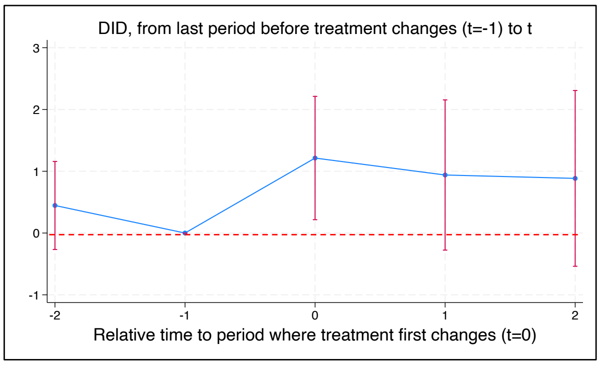
Number of income earners (in logs)



Wage income (1,0)



Savings (1,0)



Disaster relief - food and clothing (in logs)

Figure 3.9 Dynamics of income earners, wage income, savings and disaster relief

Note: Estimation results based on de Chaisemartin et al. (2023) estimator, clustered standard errors, 95% CI. Negative numbers indicate the years before the coastal El Niño in 2017, and non-negative numbers represent years after the coastal El Niño in 2017.
Source: ENAHO, 2015–2019.

In addition, I observe that support from government and private sectors increased by 385.1% after a flood disaster of this magnitude. This increase was high relative to the average amount of lost income (32.5%) and expenditure (29.9%) these households experienced.

In monetary terms, the disaster relief in the aftermath of the 2017 floods reached PEN 335.61 (USD 102.93). However, this amount is significantly lower than the income and expenditure losses faced by households between 2017 and 2019. The disaster relief only covered 22.16% of the income loss and 29.08% of the expenditure loss per capita, indicating a substantial gap between the aid provided and the actual financial impact on affected households. The limited disaster relief may hinder recovery efforts, forcing households to rely on their own savings, which exacerbates their vulnerability in the long term.

3.8 Policy recommendation

The findings highlight the significant economic impact of flooding on household incomes and expenditures per capita, as evidenced by the 2017 coastal El Niño event in northern Peru. As discussed above, support in the form of food and clothing donations falls short of fully addressing the financial strain on affected households. To enhance resilience, it is recommended that disaster insurance be introduced with coverage levels at least equivalent to the difference between the value of post-disaster donations and the expenditure loss amount (PEN 818.51 or USD 251.02), thereby ensuring a baseline of financial security.

Moreover, the damage index, which identifies districts with a higher probability of vulnerability to flood events, could be utilized by central and local governments to guide budget reallocation for post-disaster financing. This index offers an empirical foundation for understanding how local governments adjust their spending in response to natural shocks, ensuring that resources are effectively allocated to the most vulnerable areas.

3.9 Conclusions

This paper estimates the causal impact of floods on household welfare outcomes and poverty, with the coastal El Niño serving as a treatment. This phenomenon caused destructive flooding in many districts in northern Peru in 2017. Isolated extreme weather also brought flood events with less damaging effects in 2018 and 2019 within the same area. As a proxy for the local economic impact, a novel damage index is constructed using remote sensing, and combined it with five-year household panel data. I employ a diff-in-diff event study estimation strategy with heterogeneous treatment effects across groups and over time, accounting for treatments that switch on and off.

The results in this article reveal that households affected by floods experience lower income and expenditure per capita compared to households located in unaffected areas between 2015 and 2019. In addition, the findings confirm the presence of dynamic effects following the floods in 2017. There is evidence of a slow recovery process of income and expenditure, yet levels remain below those prior to the flood event. There has also been an increase in poverty, particularly among those located in urban areas, pushing more people below the poverty line.

Another key finding is that the decrease in the number of income earners within the household and the reduction in wage income are potential factors influencing the decline in income and expenditure between 2017 and 2019. Conversely, household financial assets, such as savings, and disaster relief served as ex-post coping strategies to smooth consumption expenditure among affected households. The loss of jobs, particularly in urban areas, may have resulted in fewer working members, leading to reduced income and expenditure.

However, household savings and donations of food and clothing provided to affected households alleviated the adverse impact of floods and contributed to the slight economic recovery of the affected regions.

A key lesson from this study is the importance of recognising that family savings as well as both private donations channelled through NGOs and public sector resources (from foreign and/or domestic sources) available in the aftermath of flood events, act as support measures for ex-post coping mechanisms among affected households. However, it seems that household recovery primarily depends on donations since the level of support has remained unchanged in the post-disaster period (2018–2019). This can create a cycle of dependency that hinders the development of local resilience and self-reliance. It may discourage households and communities from investing in long-term solutions to mitigate the impact of future flood events. Disaster insurance could be introduced with coverage equal to or greater than the gap between post-disaster donations and expenditure losses (PEN 818.51 or USD 251.02) to make households in affected districts more resilient when flooding of this magnitude, such as the coastal El Niño, happens again.

Finally, it is clear that affected households have not fully recovered in the aftermath of the floods. The risk of a possible post-disaster poverty trap is especially acute if households depend exclusively on external support and lack access to formal risk-coping strategies, such as microinsurance. The empirical literature has already discussed the links between disasters and poverty traps (Janzen and Carter, 2013; Hallegatte et al., 2016), suggesting that insurance payments help stabilise consumption for poor households and help protect assets for those who are relatively well-off. In that context, it is crucial to explore opportunities for developing an insurance product to cover losses caused by floods among households located in areas with high risk of such events.

References

Abrahams, A, Oram, C, Lozano-Gracia, N 2018, 'Deblurring DMSP night-time lights: a new method using Gaussian filters and frequencies of illumination', *Remote Sensing of Environment*, vol. 210, no. 1, pp. 242–258.

Bakkensen, L, Ma, L 2020, 'Sorting over flood risk and implications for policy reform', *Journal of Environmental Economics and Management*, vol. 104, 102362.

Barnston, A 2017, 'April 2017 ENSO update: Conflicting signals from the tropical Pacific Ocean', < https://www.climate.gov/news-features/blogs/enso/april-2017-enso-update-conflicting-signals-tropical-pacific-ocean>.

Becker, E 2017, 'January 2017 ENSO update: Happy new year', <www.climate.gov/news-features/blogs/enso/january-2017-enso-update-happy-new-year>.

Bertinelli, L, Strobl, E 2013, 'Quantifying the local economic growth impact of hurricane strikes: an analysis from outer space for the Caribbean', *Journal of Applied Meteorology and Climatology*, vol. 52, no. 8, pp. 1688–1697.

Bluhm, R, Krause, M 2018, 'Top lights - bright cities and their contribution to economic development', CESifo Working Paper, no. 7411.

Caetano, C, Callaway, B 2024, 'Difference-in-Differences with time-varying covariates in the parallel trends assumption', < https://arxiv.org/pdf/2406.15288>.

Callaway, B, Sant'Anna, PH 2021, 'Difference-in-differences with multiple time periods'', *Journal of Econometrics*, vol. 225, no. 2, pp. 200–230.

Carter, MR, Little, PD, Mogues, T, Negatu, W 2007, 'Poverty traps and natural disasters in Ethiopia and Hondura's, *World Development*, vol. 35, issue 5, pp. 835-856.

Centre for Research on the Epidemiology of Disasters 2018, 'Natural Disasters 2018', *CRED*, <www.cred.be/sites/default/files/CREDNaturalDisaster2018.pdf>.

Chantarat, S, Lertamphainont, S, Samphantharak, K 2016, 'Floods and farmers: evidence from the field in Thailand', PIER Discussion Paper no. 40, Puey Ungphakorn Institute for Economic Research, Bangkok.

Del Ninno, C, Dorosh, PA, Smith, LC, Roy, DK 2001, 'The 1998 floods in Bangladesh: disaster impacts, household coping strategies, and response', Research Report 122, IFPRI.

de Chaisemartin, C, D'Haultfœuille, X 2023, 'Two-way fixed effects and differences-in-differences with heterogeneous treatment effects: a survey', *The Econometrics Journal*, vol. 26, issue 3, pp. 1-30.

Elliott, R, Strobl, E, Sun, P 2015, 'The local impact of typhoons on economic activity in China: a view from outer space', *Journal of Urban Economics*, vol. 88, issue C, 50-66

ENFEN (Multisectoral Committee for the Study of El Niño) 2017, 'Informe tecnico extraordinario N°001–2017/ENFEN EL NIÑO COSTERO 2017' (in Spanish).

Elvidge, C, Baugh, K, Zhizhin, M, Hsu, FC 2013, 'Why VIIRS data are superior to DMSP for mapping night time lights', *Proceedings of the Asia-Pacific Advanced Network*, vol. 35 no, 1, pp. 62–69.

Felbermayr, G, Gröschl, J, Sanders, M, Schippers, V, Steinwachs, T 2022, 'The economic impact of weather anomalies', *World Development*, vol. 151, 105745.

Gibson, J, Olivia, S, Boe-Gibson, G, Li, C 2021, 'Which night lights data should we use in economics, and where?', *Journal of Development Economics*, vol. 149, pp. 1-12.

Gibson, J, Jiang, Y, Zhang, X, Boe-Gibson, G 2024, 'Are disaster impact estimates distorted by errors in popular night-time lights data?', *Economics of Disasters and Climate Change*, vol. 8, pp. 391–416.

Goodman-Bacon, A, 2021, 'Difference-in-differences with variation in treatment timing', *Journal of Econometrics*, vol. 225, no. 2, pp. 254–277.

Guiteras, R, Jina, A, Mobarak, AM 2015, 'Satellites, self-reports, and Submersion: exposure to floods in Bangladesh', *American Economic Review: Papers & Proceedings*, vol. 105, no. 5, pp. 232–236.

Hallegatte, S, Vogt-Schilb, A, Bangalore, M, Rozenberg, J 2016, 'Unbreakable: building the resilience of the poor in the face of natural disasters', *Climate Change and Development Series*, World Bank Group, Washington D.C.

Janzen, SA, Carter, MR 2013, ‘The impact of microinsurance on asset accumulation and human capital investments: evidence from a drought in Kenya’, *ILO Research Paper*, no. 31.

Karim, A, Noy, I 2016a. ‘Poverty and natural disasters: a regression meta-analysis’, *Review of Economics and Institutions*, vol. 7, no. 2, pp. 1–26.

Kocornik-Mina, A, McDermott, TKJ, Michaels, G, Rauch, F 2020, ‘Flooded cities’, *American Economic Journal: Applied Economics*, vol. 12, no. 2, pp. 35–66.

Lertamphainont, S, Sparrow, R 2016, ‘The economic impacts of extreme rainfall events on farming households: evidence from Thailand’, *PIER Discussion Paper*, no. 45, Puey Ungphakorn Institute for Economic Research, Bangkok.

Leon, A, Kraul, C 2017, ‘Peru’s brutal season of floods leaves 94 dead, 700,000 homeless’, *Los Angeles Times*, <http://www.latimes.com/world/mexico-americas/la-fg-peru-floods-20170328-story.html>.

Lewis, B 2017, ‘Local government form in Indonesia: tax, expenditure, and efficiency effects’, *Studies in Comparative International Development*, vol. 53, pp. 25–46.

Lewis, B, Dong, S 2025, ‘The transition to direct mayoral elections in clientelistic environments: causal public spending and service delivery effects’, *Journal of Development Economics*, vol. 172, 103380, pp. 1–18

Maccini, S, Yang, D 2009, ‘Under the weather: health, schooling, and economic consequences of early-life rainfall’, *American Economic Review 2009*, vol. 99, no. 3, pp. 1006–1026.

Mueller, V, Quisumbing, A 2011, ‘How resilient are labour markets to natural disasters? the case of the 1998 Bangladesh flood’, *Journal of Development Studies*, no. 12, pp. 1954–71.

Multisectoral Commission for the Study of the El Niño Phenomenon, 2017.

Noy, I, Nguyen, C, Patel, Pooja 2020, ‘Floods and spillovers: households after the 2011 great flood in Thai’, *Economic Development and Cultural Change*, vol. 69, no. 2, pp. 829–868.

OCHA 2017, ‘Peru: Rainy season – Situation Report No. 10 (as of 10 May 2017)’, *UN Country Team in Peru,* <https://reliefweb.int/report/peru/peru-rainy-season-situation-report-no-10-10-may-2017>.

-- 2017, ‘Peru: Rainy season – Situation Report No. 12 (as of 27 June 2017)’, *UN Country Team in Peru*, 2024 <https://reliefweb.int/report/peru/peru-rainy-season-situation-report-no-12-27-june-2017>.

Parida, Y, Dash, DP, Bhardwaj, P, Chowdhury, JR 2018, ‘Effects of drought and flood on farmer suicides in Indian states: an empirical analysis’, *Economics of Disasters and Climate Change 2*, no. 2, pp. 159–80.

Poaponsakorn, N, Meethom, P, Pantakua, K 2015, ‘The Impact of the 2011 floods, and flood management on Thai households’, *In Resilience and Recovery in Asian Disasters*, pp. 75-104.

Puspito Sari, DA, Ridhani, C 2023, ‘Exploring people’s reasons of living in disaster-prone area and promoting disaster risk reduction in urban planning’, *Advances in Engineering Research*, AER 215, pp. 126–138.

Rambachan, A, Roth, J 2023, ‘A more credible approach to parallel trends’, *Review of Economic Studies*, vol. 90, no. 5, pp. 2555–2591.

Ramirez, I, Briones, F 2017, ‘Understanding the El Niño Costero of 2017: the definition problem and challenges of climate forecasting and disaster responses’, *International Journal of Disaster Risk Science*, vol. 8, pp. 489–492.

Rodriguez-Oreggia, E, De La Fuente, A, De La Torre, R, Moreno, HA (2013), ‘Natural disasters, human development and poverty at the municipal level in Mexico’, *The Journal of Development Studies*, vol. 49, issue 3, pp. 442–455.

Roth, J 2022, ‘Pretest with caution: event-study estimates after testing for parallel trends’, *American Economic Review: Insights*, vol. 4, no. 3, pp. 305-322.

Roth, J, Sant’Anna, PHC 2023, ‘Efficient estimation for staggered rollout designs’, *Journal of Political Economy Microeconomics*, vol. 1, no. 4, pp. 669–709.

Roth, J, Sant'Anna, PHC, Bilinski, A, Poe, J 2023, 'What's trending in difference-in-differences? a synthesis of the recent econometrics literature', *Journal of Econometrics*, vol. 235, pp. 2218-2244.

Sardon, H, Lavado-Casimiro, W, Felipe, O 2022, 'Inventario de datos de eventos de inundaciones del Peru. Estudio Final', *Servicio Nacional de Meteorologia e Hidrologia del Peru - SENAMHI*, <https://repositorio.senamhi.gob.pe/handle/20.500.12542/1786>.

Skoufias, E, Strobl, E, Tveit, TB 2020, 'Flood and tsunami damage indices based on remotely sensed data: an application to Indonesia', *Natural Hazards Review*, vol. 21, issue 4.

Smith, MP, Ricker, M, Prütz, R, Anand, M, Lehner, B, Flörke, M, Wimmer, F, Mann, H, Weller, D, Mucke, P, Pfliegner, K 2021, 'Global assessment of current and future river flooding and the role of nature-based solutions for risk management', *The Nature Conservancy*, Berlin.

Sun, L, Abraham, S 2021, 'Estimating dynamic treatment effects in event studies with heterogeneous treatment effects', *Journal of Econometrics*, vol. 225, no. 2, pp. 175–199.

Takahashi, K, Martínez, AG 2017, 'The very strong coastal El Niño in 1925 in the far-eastern Pacific', *Climate Dynamics*, vol. 52, pp. 7389–7415.

Tyndall Centre for Climate Change Research 2004, 'New indicators of vulnerability and adaptive capacity', *Global Forest Watch Tyndall Centre*, no. 7.

Venkateswaran, K, MacClune, K, Enriquez, MF 2017, 'Learning from El Niño Costero 2017: opportunities for building resilience in Peru', *ISET International and the Zurich Flood Resilience Alliance*.

World Bank 2008, 'Project information document (PID): concept stage – Jakarta Urgent Flood Mitigation Project (Report No. AB4043)', *East Asia and Pacific Region*. Government of Indonesia.

World Bank Group 2013, 'Why resilience matters – the poverty impacts of disasters', <https://elibrary.worldbank.org/doi/abs/10.1596/1813-9450-6699>.

Yonson, R 2018, 'Floods and pestilence: diseases in Philippine urban areas', *Economics of Disasters and Climate Change 2*, no. 2, pp. 107–35.

Appendix

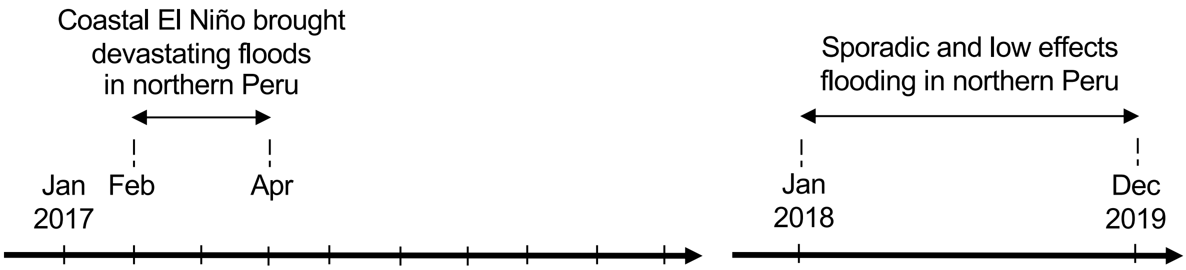


Figure 3.A.1 Timeline of the 2017 coastal El Niño and subsequent flood events

Source: Author's illustration

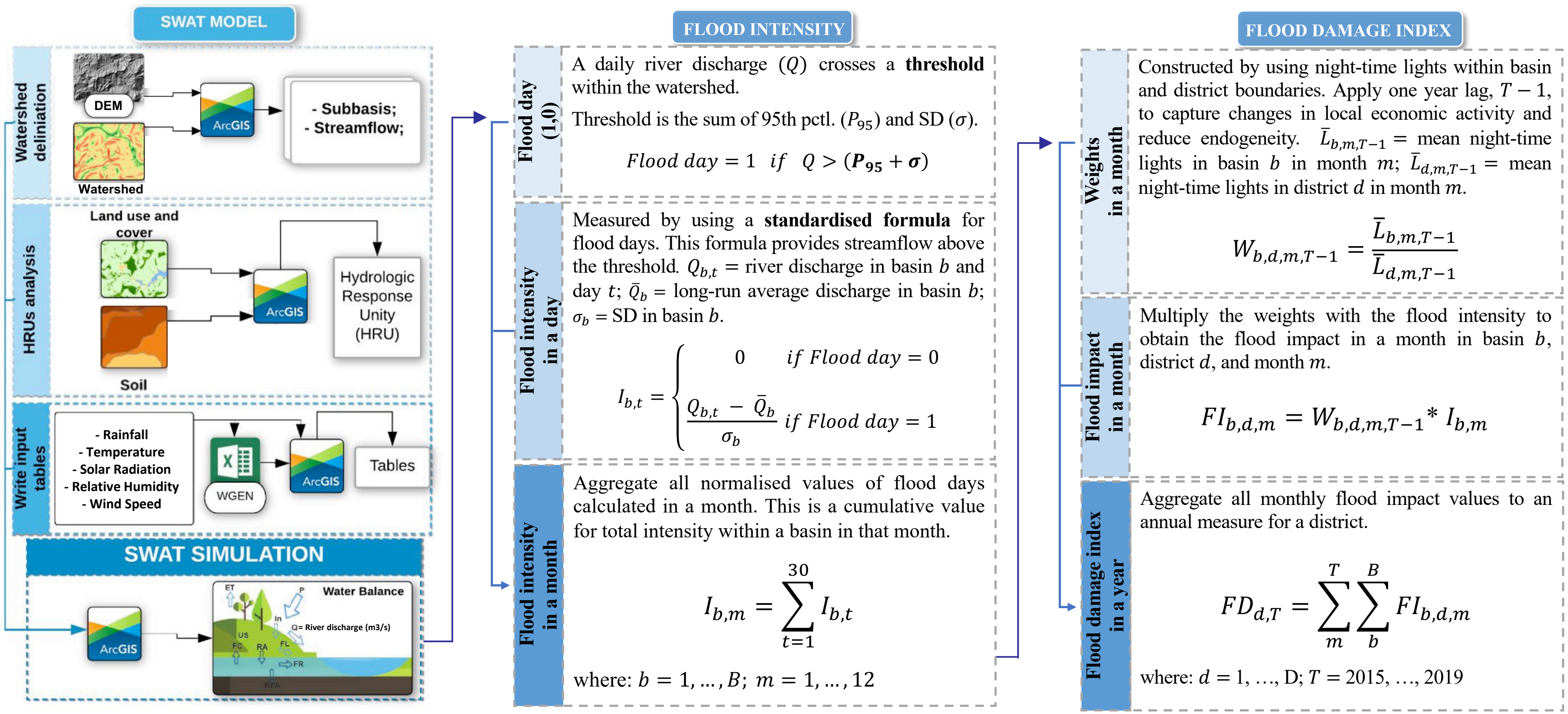


Figure 3.A.2 Flowchart of the procedure used in the hydrological simulation by ArcSWAT, and construction of the flood damage index

Source: Author's illustration

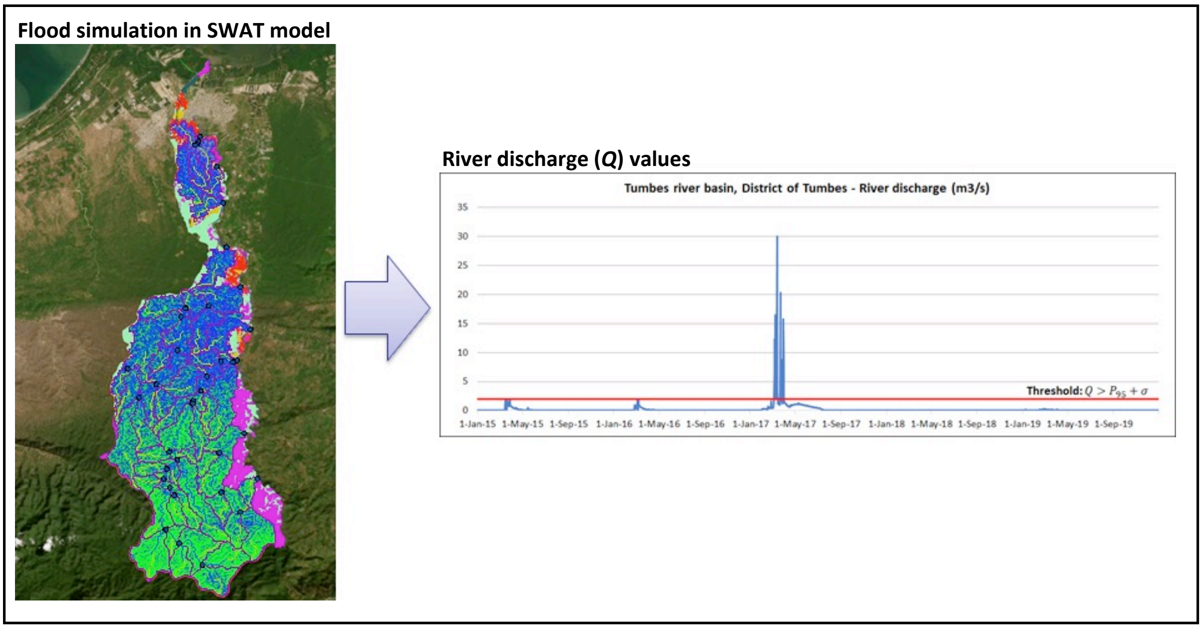


Figure 3.A.3 Watershed study design for Tumbes river basin

Note: Watershed characteristics are shown including DEM, Watershed, LULC, DSMW, weather data.
Source: Author’s illustration

Table 3.A.1 Exogeneity tests of binary switching treatment and flood damage index

Dependent variable:	2015–2017 Binary switching treatment		Flood damage index	
Annual net HH income per capita (in logs)	0.002		-9.577	
	(0.023)		(8.532)	
Annual net HH expenditure per capita (in logs)	-0.052		3.776	
	(0.049)		(12.398)	
Poverty (1,0)	-0.011		-24.966	*
	(0.033)		(12.900)	
Urban (1,0)	0.012		-19.728	
	(0.059)		(27.692)	
Wage income (1,0)	0.038		16.791	
	(0.031)		(14.851)	
Household savings (1,0)	0.111	***	15.519	
	(0.040)		(11.681)	
Disaster relief - food and clothing (in logs)	0.001		4.622	**
	(0.005)		(2.294)	
Sewage (1,0)	-0.041		-41.741	*
	(0.054)		(21.261)	
Water (1,0)	0.035		40.384	
	(0.059)		(26.469)	
Electricity (1,0)	0.027		-60.898	
	(0.100)		50.903	
Concrete roof (1,0)	-0.002		-15.912	
	(0.041)		(18.344)	
Female household head (1,0)	0.006		-8.048	
	(0.041)		(10.293)	
Household size (in logs)	-0.034		23.406	
	(0.032)		(14.454)	
Age household head (in years)	0.000		-0.444	
	(0.001)		(0.410)	
Years of schooling household head (in years)	0.005		0.289	
	(0.005)		(1.103)	
Income earners (in logs)	-0.056	*	-4.494	
	(0.032)		(10.898)	
Own house (1,0)	-0.037		3.837	
	(0.043)		(12.703)	
Public health insurance (1,0)	0.007		-11.934	
	(0.024)		(8.886)	
N	767		767	

Columns 1 and 2 report the results from regressing the treatment variable (binary and continuous) on pre-determined outcomes and exogenous variables using data from 2015 to 2017. Estimation is by OLS using both the matching weights and household sampling weights, and controlling for distance to the nearest river (in logs) and year fixed effects. Standard errors are clustered at the district level. ∗∗∗p < 0.01, ∗∗p < 0.05, ∗p < 0.1. Source: ENAHO, 2015–2019.

Table 3.A.2 Exogeneity of binary switching treatment and flood damage index for the subsample of "never treated" districts

	Period 2016–2019	
Dependent variable:	Binary switching treatment	Flood damage index
Annual net HH income per capita (in logs)	0.000	-0.002
	(0.001)	(0.007)
Annual net HH expenditure per capita (in logs)	0.004	0.038
	(0.004)	(0.042)
Poverty (1,0)	0.002	0.019
	(0.003)	(0.026)
Urban (1,0)	-0.004	-0.043
	(0.005)	(0.047)
Wage income (1,0)	0.000	-0.004
	(0.001)	(0.012)
Household savings (1,0)	-0.006	-0.059
	(0.006)	(0.062)
Disaster relief - food and clothing (in logs)	0.000	0.004
	(0.000)	(0.004)
Sewage (1,0)	0.000	0.002
	(0.003)	(0.026)
Water (1,0)	0.001	0.012
	(0.002)	(0.020)
Electricity (1,0)	-0.002	-0.021
	(0.003)	(0.026)
Concrete roof (1,0)	0.004	0.037
	(0.004)	(0.041)
Female household head (1,0)	0.010	0.103
	(0.011)	(0.107)
Household size (in logs)	-0.007	-0.070
	(0.007)	(0.072)
Age household head (in years)	0.000	-0.001
	(0.000)	(0.001)
Years of schooling household head (in years)	0.000	-0.002
	(0.000)	(0.003)
Income earners (in logs)	0.006	0.060
	(0.006)	(0.061)
Own house (1,0)	-0.008	-0.080
	(0.009)	(0.085)
Public health insurance (1,0)	-0.005	-0.053
	(0.006)	(0.056)
N	456	456

Columns 1 and 2 show the results of regressing the treatment variable (binary and continuous) on predetermined outcomes and exogenous variables for the subsample of "never treated" districts (i.e., for those districts that have not experienced flooding) from 2016 to 2019. Estimation is by OLS using both the matching weights and household sampling weights, and controlling for distance to the nearest river (in logs) and year fixed effects. Standard errors are clustered at the district level. ∗∗∗$p < 0.01$, ∗∗$p < 0.05$, ∗$p < 0.1$. Source: ENAHO, 2015–2019.

Table 3.A.3 "Switchers" districts in northern Peru

SWITCHING UNITS Districts:	Year				
	2015	2016	2017	2018	2019
Morrope	0	0	2,472.90	10.86	0
Frias	0	0	1,761.32	0	0
Pacaipampa	0	0	1,092.62	0	0
Chulucanas	0	0	828.01	0	0
Tambo Grande	0	0	744.54	0	1.77
Las Lomas	0	0	679.90	0	0
La Brea	0	0	497.23	0	4.87
Olmos	0	0	343.36	0	3.17
Tumbes	0	0	325.37	0	0
El Prado	0	0	252.53	53.90	0
Cura Mori	0	0	205.83	0	0
Jayanca	0	0	187.86	10.06	0
Sondorillo	0	0	174.23	0	1.62
Pomalca	0	0	158.19	11.50	0
San Andres de Cutervo	0	0	130.93	65.79	0
Pampas de Hospital	0	0	112.89	0	0
La Cruz	0	0	91.64	0	0
Viru	0	0	68.44	128.92	0
Pacora	0	0	64.30	49.57	0
Julcan	0	0	52.27	49.33	0
Castilla	0	0	50.25	0	0
Guadalupito	0	0	44.71	0	0
Chalamarca	0	0	36.66	63.25	0
Chota	0	0	26.17	24.41	0
Sullana	0	0	21.58	0	0
Pimentel	0	0	18.64	0	0
Piura	0	0	12.06	0	0.34
Guadalupe	0	0	11.77	12.34	0
Bambamarca	0	0	11.44	46.45	0
Tuman	0	0	10.60	17.82	0
Encañada	0	0	8.42	41.36	0
Trujillo	0	0	5.91	9.86	0
Veintiseis de Octubre	0	0	4.64	0	0.11
Reque	0	0	1.38	9.06	0

Note: The table shows the switching units (districts) as treatment (flooding) switches on and off over time (2017–2019) and across units (districts), with treated units (districts) receiving treatment (floods) in the same year (2017). The devastating coastal El Niño floods in 2017 affected all districts, while sporadic and less severe floods in 2018 impacted 16 districts, and only six districts experienced flood events in 2019.
Source: ENAHO, 2015–2019.

Table 3.A.4 "Always treated" districts in northern Peru

TREATED UNITS Districts:	Year				
	2015	2016	2017	2018	2019
Aguas Verdes	0	0	561.46	3.67	6.65
La Victoria	0	0	39.73	5.18	0.48
Llacanora	0	0	21.84	44.65	3.02
Huamachuco	0	0	14.60	55.24	4.34
Chiclayo	0	0	11.30	1.55	0.14
Cajamarca	0	0	4.29	12.75	0.85

Note: The table shows the always units (districts) as treatment (floods) happened over time (2017–2019) and across units (districts), with treated units (districts) receiving treatment (floods) in the same year (2017). The devastating coastal El Niño floods in 2017 affected all districts, while sporadic and less severe floods in 2018 impacted 16 districts, and only six districts experienced flood events in 2019.
Source: ENAHO, 2015–2019.

Table 3.A.5 "Never treated" districts in northern Peru

CONTROL UNITS Districts:	Year				
	2015	2016	2017	2018	2019
Bellavista	0	0	0	0	0
Bolivar	0	0	0	0	0
Cajabamba	0	0	0	0	0
Canchaque	0	0	0	0	0
Canoas de Punta Sal	0	0	0	0	0
Chepen	0	0	0	0	0
Cochorco	0	0	0	0	0
Corrales	0	0	0	0	0
Domingo de la Capilla	0	0	0	0	0
El Porvenir	0	0	0	0	0
Ferreñafe	0	0	0	0	0
Jose Leonardo Ortiz	0	0	0	0	0
La Arena	0	0	0	0	0
La Esperanza	0	0	0	0	0
Moche	0	0	0	0	0
Salpo	0	0	0	0	0
Zarumilla	0	0	0	0	0

Note: The table shows the control units (districts). These units are never treated during the analysis period. They serve as the comparison group to assess the impact of the treatment.
Source: ENAHO, 2015–2019.

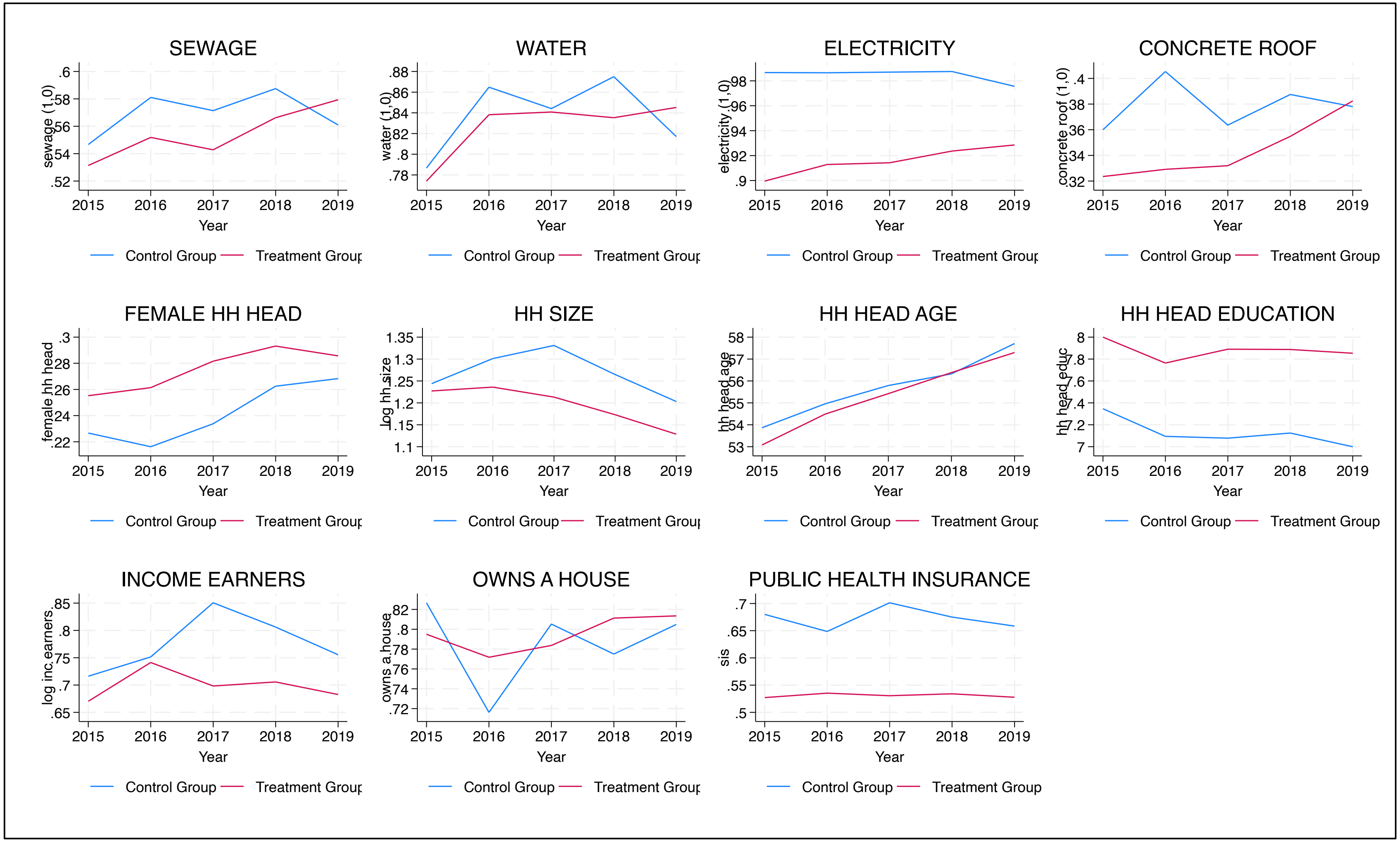


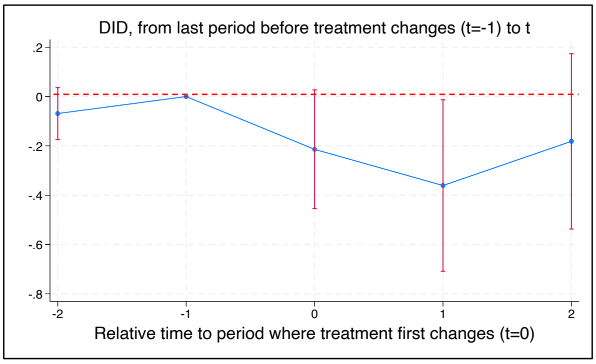
Figure 3.A.4 Average covariates over time for treated and control groups

Note: Plots for average covariates that are included in the model, which evolve in a non-parallel fashion across the treatment and control groups.
Source: Author's illustration.

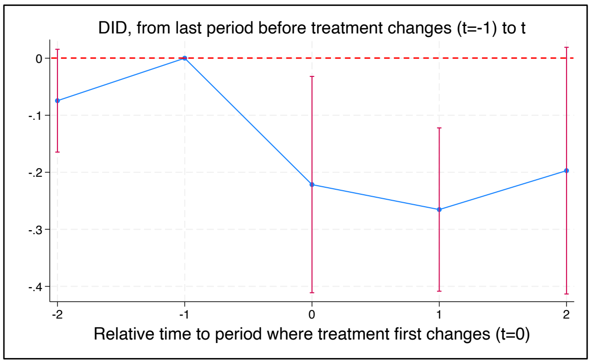
Table 3.A.6 Impact of floods on household welfare and poverty using weights based on 2013 VIIRS night-time lights

Dependent variable:	**Income per capita (in logs)**		**Expenditure per capita (in logs)**		**Poverty (1,0)**	
A: Flood damage level						
Effect -1 (year 2015)	-0.069		-0.075		0.065	
	(0.054)		(0.046)		(0.043)	
Effect 0 (year 2017)	-0.214	*	-0.222	**	0.186	**
	(0.123)		(0.097)		(0.067)	
Effect 1 (year 2018)	-0.361	**	-0.265	***	0.315	***
	(0.177)		(0.073)		(0.092)	
Effect 2 (year 2019)	-0.182		-0.197	*	0.327	***
	(0.181)		(0.110)		(0.089)	
ATT (flooded district=index)	**-0.007**	*	**-0.007**	***	**0.008**	***
	(0.004)		**(0.002)**		**(0.002)**	
District FE	Yes		Yes		Yes	
Year FE	Yes		Yes		Yes	
Cluster (district)	Yes		Yes		Yes	
Controls	Yes		Yes		Yes	
N	913		913		913	

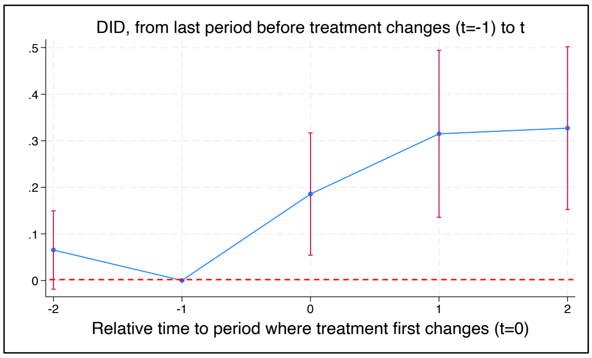
The dependent variables are listed across the top row. Income and expenditure per capita are measured as the log of net annual household income and expenditure per member, respectively. Poverty is a binary variable that takes on a value of one if the households is below the poverty line, else zero. Estimation results are based on the de Chaisemartin et al. (2023) estimator, using a flood damage index constructed from 2013 VIIRS night-time lights data. The model controls for dwelling, household, and geographic characteristics, as well as district and year fixed effects. Standard errors are clustered at the district level. ∗∗∗ $p < 0.01$, ∗∗ $p < 0.05$, ∗ $p < 0.1$.
Source: ENAHO, 2015–2019.



Income per capita (in logs)



Expenditure per capita (in logs)



Poverty (1,0)

Figure 3.A.5 Event study plots for impact of floods on household welfare and poverty using weights based on 2013 night-time lights

Note: Estimation results based on de Chaisemartin et al. (2022) estimator, clustered standard errors, 95% CI. Negative numbers indicate the years before the coastal El Niño in 2017, and non-negative numbers represent years after the coastal El Niño in 2017.
Source: ENAHO, 2015–2019.

Table 3.A.7 Impact of floods on household savings

Dependent variable:	Household savings (1,0)		
	Urban		**Rural**
A: Binary treatment			
ATT (flooded district=1)	-0.551	**	-0.127
	(0.274)		(7.095)
B: Flood damage level			
ATT (flooded district=index)	-0.007	**	-0.001
	(0.004)		(0.039)
District FE	Yes		Yes
Year FE	Yes		Yes
Cluster (district)	Yes		Yes
Controls	Yes		Yes
N	539		374

The dependent variable is listed in the top row. Savings in urban areas equals one if the household head has a bank account. Savings in rural areas equals zero if the household head has a bank account in the rural area. Estimation results are based on the de Chaisemartin et al. (2023) estimator, controlling for dwelling, household and geographic characteristics, as well as district and year fixed effects. Standard errors are clustered at the district level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.
Source: ENAHO, 2015–2019.

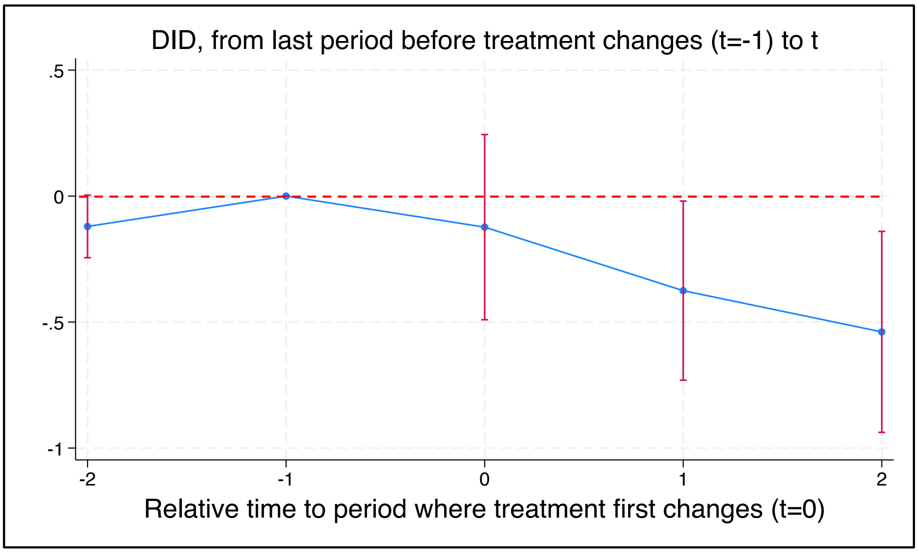


Figure 3.A.6 Event study plot for impact of flood shocks on household savings (1,0) in urban areas

Note: Estimation results based on de Chaisemartin et al. (2023) estimator, clustered standard errors, 95% CI. Negative numbers indicate the years before the coastal El Niño in 2017, and non-negative numbers represent years after the coastal El Niño in 2017.
Source: ENAHO, 2015–2019.

Chapter 4
Preferences, trust, and index-based flood insurance: Evidence from Peru

4.1 Introduction

Floods are the most frequent and among the costliest weather-related disasters worldwide, with nearly one in four people—approximately 1.81 billion—directly exposed to flood risks. Of those at risk, 89% live in low- and middle-income countries, where the capacity to respond to such disasters is often limited (Rentschler et al., 2022). As global warming intensifies, the frequency and severity of floods are expected to rise, disproportionately affecting densely populated urban areas, many of which have expanded into flood-prone zones (Brunner et al., 2021). The economic impacts of these disasters are staggering, with damages in 2023 alone estimated at US$20.4 billion (CRED, 2023). The increasing risks posed by climate change, combined with rapid urbanisation, highlight the need for effective flood risk management, particularly for vulnerable populations living in high-risk areas.

In this study, I examine the role of economic preferences—specifically time and risk preferences, and price—along with trust in institutions, in explaining households' decisions to adopt a hypothetical index-based flood insurance product in an urban context. This product, also known as parametric insurance, provides compensation for predicted losses based on a predefined index linked to specific disaster events. The focus is on Lima, Peru, a megacity vulnerable to flooding. The central research question is: How do economic preferences and trust influence the demand for index-based flood insurance in vulnerable communities?

To investigate these factors, a lab-in-field experiment is conducted in Lurigancho–Chosica, a flood-prone district in Lima, offering a hypothetical index insurance product based on rainfall-related indices correlated with flood events that cause household losses. The experiment includes games designed to elicit two economic preferences: a time preference game, where households choose between an immediate, smaller reward (i.e., strong time preference) and a larger, delayed reward (i.e., weak time preference), grouped by three choice sets; and a risk preference game, where households select one of six investment options with varying levels of risk and payoff, allowing for an estimation of their risk aversion. I experimentally vary the order in which each choice set is presented to the household, allowing for the randomisation of their intertemporal decisions. The insurance demand is captured by using a multiple price list, where households are randomly assigned one of five monthly premium rates. In addition, a survey assessed households' trust in local government policy development, flood management plans, and the insurance system to understand how trust influences insurance uptake.

Index-based insurance has gained attention in rural areas of developing countries as a tool for managing weather-related risks, with empirical studies highlighting its role in reducing asset sales, smoothing consumption, and increasing access to credit (Kazianga and Udry, 2006; Janzen and Carter, 2013). Despite these benefits, the adoption of index-based insurance remains low, even in the face of subsidised schemes and experimental interventions designed to reduce barriers to uptake (Giné et al., 2008; Binswanger-Mkhize, 2012; Cole et al., 2013). While prior studies have focused on factors such as price, financial literacy, and transaction costs (Giné et al., 2008), there has been limited attention to how behavioural factors—such as time and risk preferences—and trust in institutions influence insurance demand, especially in urban areas. This gap is particularly significant because index-based insurance, by its nature, involves a trade-off between an upfront premium payment and uncertain future payouts, making behavioural factors crucial in understanding uptake.

In addition, multi-peril index-based insurance products covering flood damage are available for smallholder farmers in developing countries; however, there is currently no index-based insurance product focused solely on flood-related losses (Malik and Amarnath, 2021). By and large, index-based insurance has yet to be applied in urban settings, where it could help mitigate the financial risks faced by vulnerable communities exposed to frequent flooding.

The empirical results suggest a positive relationship between highly weak time preference and demand for insurance, illustrating the importance of timing in premium payments when trying to understand index-based insurance decisions. I also find that extreme risk aversion discourages insurance uptake, supporting the idea that risk-averse households are sceptical of the product benefits as a risk management tool. The intuition behind this is the lack of trust in the government's ability to handle floods or manage insurance programs effectively. Nevertheless, the findings suggest that higher trust in government can facilitate insurance uptake, as households are more likely to participate when they believe in the reliability and transparency of the compensation mechanisms. These steps could increase household confidence and encourage insurance participation.

This paper contributes to the literature in three key ways. First, it shows evidence of how economic preferences and their interaction affect the adoption of a hypothetical index-based insurance product. Limited research has studied the effects related to time and risk preferences, and price on the decision to insure against weather-related impacts (Baillon et al., 2022; Kundu 2022; Shin et al., 2022; Hill et al., 2016). By conducting a lab-in-field experiment, this study, on one hand, highlights the importance of households' time preferences, particularly their willingness to pay for insurance upfront, significantly influence their choice to purchase insurance. This aligns with existing literature that suggests liquidity constraints and the timing of payments play a critical role in insurance uptake (Belissa et al., 2019; Casaburi and Willis, 2018; Liu et al., 2020; Baillon et al., 2022). On the other hand, this research draws the attention of the demand response to changes in risk preferences, when households must weigh uncertain future outcomes, such as potential income losses and presence of basis risk—inability to perfectly represent the losses of an insured household—due to flood events.

Second, the paper explores how the level of trust in local government and institutions interacts with economic preferences to shape insurance demand. Trust has been shown to be a critical factor in the adoption of insurance, yet its role in urban settings, particularly regarding government-managed flood insurance programs, has received less attention. The weather insurance literature has presented findings on the importance of trust in uptake, showing the expected association but usually with no statistical power. Moreover, trust in institutions, including government, has been shown to encourage insurance purchases to increase disaster preparedness (Barnes et al., 2020; Cao et al., 2020; Ejeta et al., 2018; Soane et al., 2010), particularly in vulnerable areas of developing countries, such as rural settings where communities are exposed to weather shocks. Conversely, low trust in government can heighten vulnerability, especially where government presence is limited (Ullah et al., 2021).

Third, this study is one of the first to introduce a hypothetical index-based flood insurance product for the urban context of Lima, Peru. This megacity, one of the biggest and most populated in Latin America, is particularly vulnerable to flood-related disasters, making it a critical case study. The district of Lurigancho–Chosica is in constant flood risk due to factors such as: i) high rainfall intensity, typically between January and April; ii) presence of informal settlements along arid watercourses prone to flash flooding; and iii) inadequate public infrastructure for flood mitigation and response. Consequently, this case study provides a valuable setting to examine the potential demand for flood insurance as a tool for enhancing household resilience against recurrent climate-induced shocks.

The paper is organised as follows: Section 4.2 provides background information on the index-based insurance product and its application in the Lurigancho–Chosica district. Section 4.3 details the lab-in-field experiment and the methodology used to measure economic preferences and trust. Section 4.4 outlines the identification strategy, while Section 4.5 presents results and extended analysis. Finally, Sections 4.6 and 4.7 present policy implications and concluding remarks, respectively.

4.2 Background

Lima, a megacity with a population of 10.15 million (INEI Bulletin, 2023), is prone to annual flooding between January and April. The city's eastern side has experienced significant flood events in 2012, 2015, 2017, and 2023, placing vulnerable populations and areas at high risk. In 2017, intense rainfall triggered severe floods that affected several parts of the city, particularly areas along the Rimac River, which flows from east to west, and the surrounding Andes hillsides. Water descending rapidly from these highlands often carries debris and sediment, exacerbating flooding in the lower-lying urban districts. One of the hardest-hit areas is Lurigancho–Chosica, where urban settlements near the watercourse remain vulnerable to annual flooding.

The National Institute of Civil Defence (INDECI) has identified flood-risk hotspots in Chosica, where 73,919 people (28.84% of the total population of 256,294) are exposed to severe flooding (National Water Authority, 2017). Additionally, the district faces challenges related to poverty and desertification (ACT Alliance, 2012). In the urban communities of Nicolás de Piérola, Sierra Limeña, and San Antonio de Pedregal, approximately 70% of households are classified as poor and vulnerable, according to the sample. Flooding has caused considerable financial losses and property damage to these populations over the last two decades.

In this context, the central and local governments have undertaken various efforts to mitigate and protect vulnerable populations residing near watercourses and hillsides from the adverse effects of such weather events. Among these measures are retention walls, installed in 2017, designed to channel water flow during intense rainfall or floods. Dynamic barriers made of high-strength steel were placed at critical points upstream (see Figure 4.1) in 2016 to prevent rock and debris flows from reaching urban areas downstream. Moreover, the government has been engaged in reforestation efforts to prevent soil erosion and has collaborated with national agencies such as the National Water Authority (ANA) to remove large quantities of accumulated debris from these areas. Despite these interventions, the population remains at risk, and many continue to be affected by the recurring floods and landslides. These combined efforts, while helpful, have not fully eliminated the flood threats in those urban communities.

To explore opportunities for developing a weather risk transfer product, a lab-in-field experiment was conducted and households were offered a hypothetical index-based flood insurance product designed to cover financial losses and provide liquidity to smooth consumption in the urban communities of Nicolás de Piérola, Sierra Limeña, and San Antonio de Pedregal. When households were visited by enumerators, they were informed that the local government of Lurigancho–Chosica plans to develop and launch index-based insurance for all who are interested in the insurance scheme. To ensure respondents clearly understood the product being evaluated, the questionnaire contained a standardised script describing the hypothetical index insurance and how it differs from traditional insurance, the role of the local weather station in triggering payouts, the rainfall threshold associated with severe flooding in their area, and the fixed compensation amount if the threshold was exceeded. Respondents were presented with an example based on their valley location, including the specific rainfall threshold in millimetres and a historical reference to the 2017 floods. The exact text read to participants is provided in Module K01, survey questionnaire section – Household survey on the impact of floods in Lima.

The monthly premium is PEN 227.27 (USD 61.61) per household[1], compared to an average monthly household expenditure of PEN 1,337.43 (USD 362.55) according to the ENAHO survey in 2023. This insurance scheme will only be implemented if a large number of households are willing to participate.

[1] The premium is determined by the insurance compensation and the probability of a flood event. I multiply the indemnity of PEN 10,000 by the flood probability, which is 27%. This probability is based on the average interval between flood years (2012, 2015, 2017, and 2023). The resulting amount is then divided by 12, which gives a monthly premium of PEN 227.27 for the index-based insurance product.

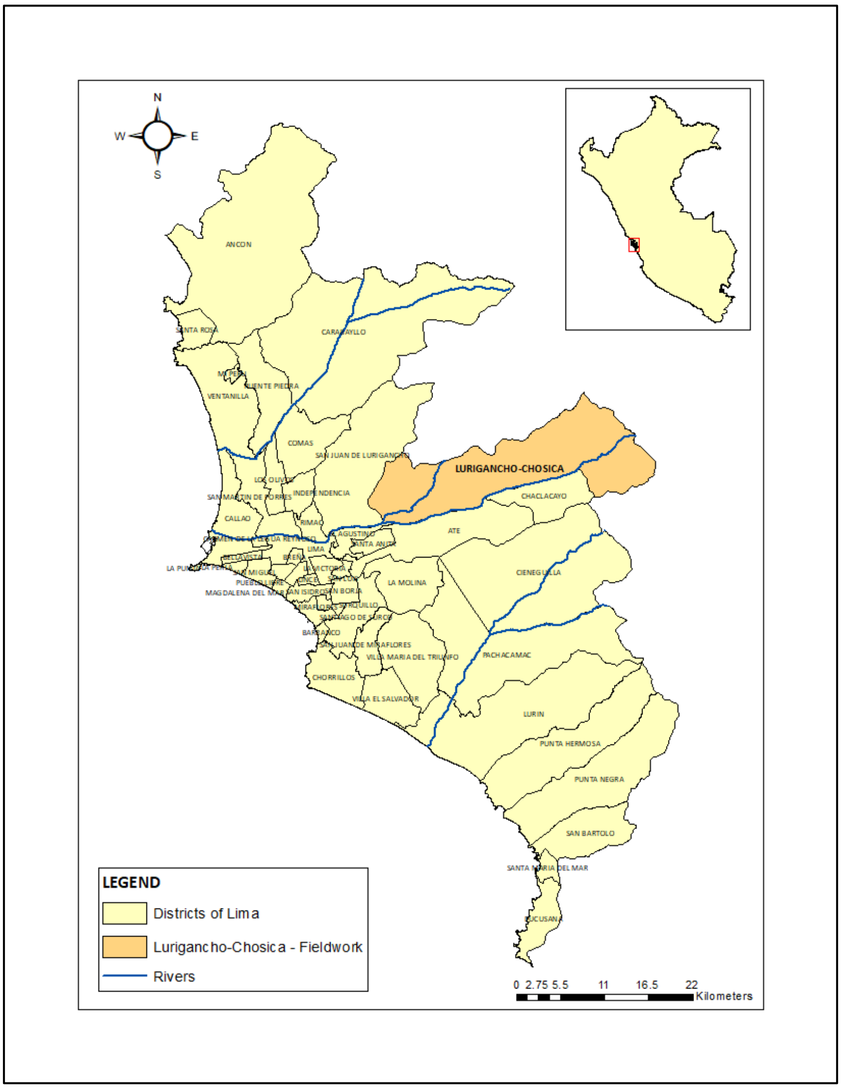


Figure 4.1 District of Lurigancho–Chosica vulnerable to flooding

Note: This figure shows the district of Lurigancho–Chosica, where vulnerable households are at high flood risk, and where the lab-in-field experiment was conducted.
Source: Author's design, 2023.

If the level of rainfall recorded at the weather station located in the district of Lurigancho-Chosica reaches an accumulated 36.4 millimetres or more[2], all households that are policyholders will be compensated PEN 10,000 (USD 2,711).

2 Using historical data from the National Service of Meteorology and Hydrology weather station located in the district of Lurigancho-Chosica, I determine that the amount of accumulated rainfall in mid-March 2017 caused severe floods in the areas near the watercourse and hillsides in the three urban communities.

4.3 Data

4.3.1 Lab-in-field survey

To test the implications of economic preferences on insurance demand, I conducted a lab-in-field experiment for a month between May and June in 2023. The sample includes 404 households living in the urban communities of Nicolas de Piérola, Sierra Limeña, and San Antonio de Pedregal, which are nestled within the district of Chosica, east of the capital, Lima, in Peru[3].

Due to varying flood risks, I classified areas within the urban communities into three categories: 1) flood-free; 2) flood-prone hillsides; and 3) rainy season riverbeds. Areas classified as flood-prone hillsides experience rapid flooding and erosion even with small amounts of localized rainfall. In contrast, areas in rainy season riverbeds are prone to flooding during heavy rainfall, which brings water downstream from the valley bottom along with significant sediment loads.

To implement the study, I partnered with a local research group from the National University of Engineering in Lima. They provided us with a fieldwork supervisor and a collection of Peru's official information on: i) flood risk maps from the National Centre for Estimation, Prevention and Reduction of Disaster Risk (CENEPRED); ii) housing listings from the Municipality of Chosica; and 3) land titles from the Ministry of Housing. They also facilitated contact with the urban community leader of Nicolás de Piérola to obtain permission to conduct fieldwork there. During the pre-fieldwork phase, I was also able to contact the community leader of San Antonio de Pedregal to secure permission to enter the urban community.

The lab-in-field experiment elicited (1) time preferences, (2) risk preferences, and (3) premium rates. Experimental games in (1) and (3) have the order of choices randomised at the respondent level. For (2), a separate game was designed to estimate the degree of risk aversion. All experiments were hypothetically incentivised, but only (2) involved actual payments at the end of the game, while payments related to risk preferences were made immediately. All sampled household were assigned to these three games at the same time so that they were equally exposed to them.

---

[3] See Appendix Section A for more details about the calculation of the sample size for the lab-in-the-field experiment.

A trust game was not included in the experiment; however, trust was examined through targeted survey questions. For example, respondents were asked about their level of trust in various potential index insurance providers. The survey also captured perceptions of the local government's capacity and integrity in managing flood insurance, its effectiveness in responding to floods, and its role in delivering infrastructure, education, and health programs.

Figure 4.2 displays the locations of survey participants, geodynamic barriers, and floodwater paths in the district of Chosica. Households in the sample are coloured dark umber, and the barriers are coloured green. According to flood risk exposure, rainy season riverbeds are coloured light blue, and flood-prone hillsides are coloured light red.

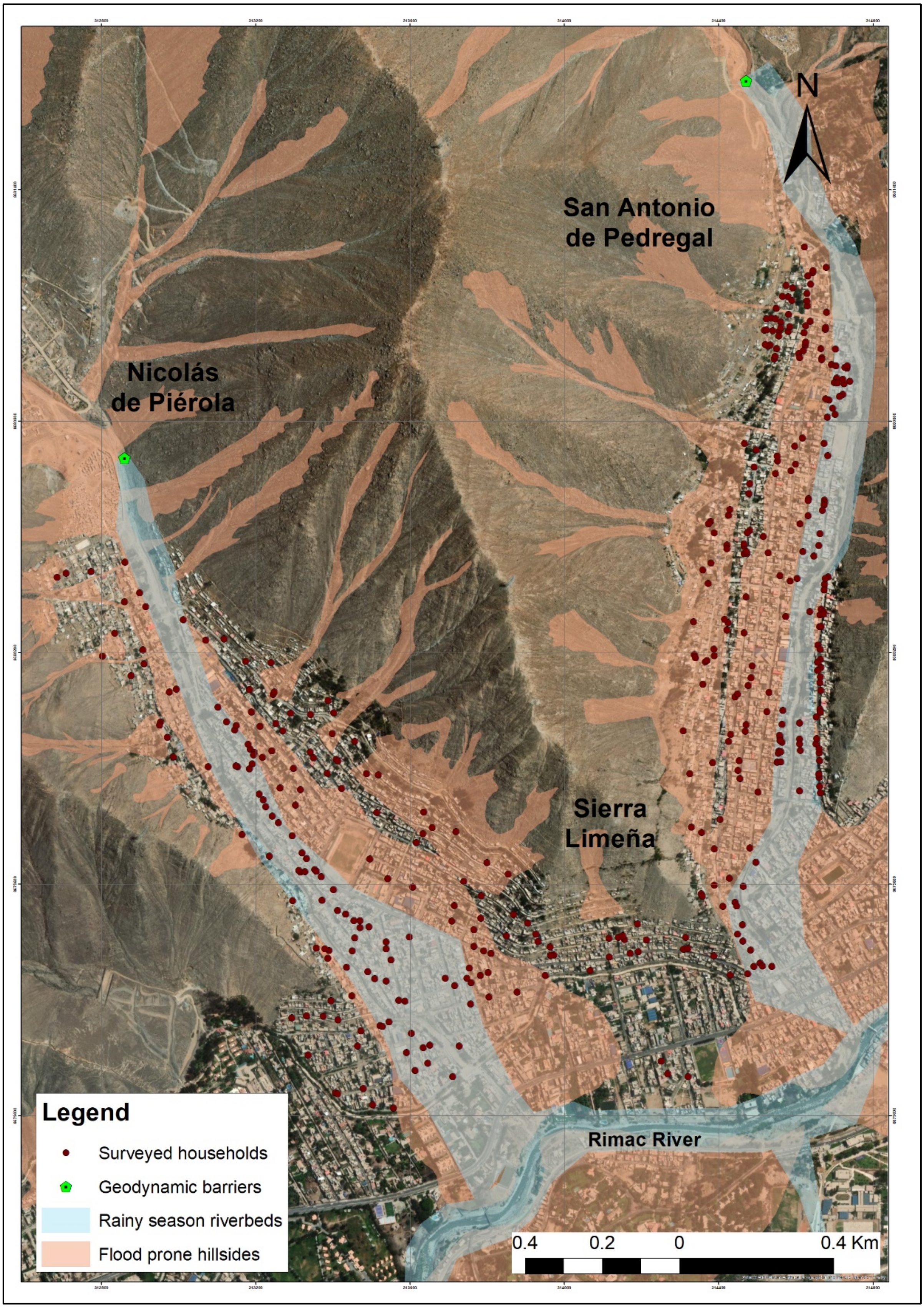


Figure 4.2 Location of households and geodynamic barriers, and flood water paths

Note: This figure shows the urban communities of the study area, the households who participated in the survey, the location of the geodynamic barriers and flood risk areas.
Source: Author's design, 2023.

Economic preference games

Time preferences — I used a simple choice list approach, where respondents faced a trade-off between a sooner (but smaller) reward and a later (but larger) reward (Bauer, Chytilova, and Morduch 2012; Almås et al. 2016; Chowdhury 2022). The choice lists were kept as simple as possible to ensure participants understood the choice options – all payments involve hypothetical money; however, respondents were asked to make their decision as if an actual payment was at stake. Table 4.1 presents the 12 choices participants were required to make. I designed three sets of choices: in the first two sets, the earliest payment was scheduled between tomorrow, one month, and one year, while the later payment was scheduled between three months, four months, and one year and three months. The first two sets had a delay of three months between the earliest and later payments, whereas the third set had a delay of one year.

Table 4.1 Time preferences

	Option 1			Option 2		
Choice set 1	PEN 5.00	USD 1.36	tomorrow	PEN 5.30	USD 1.44	3 months
	PEN 5.00	USD 1.36	tomorrow	PEN 6.00	USD 1.63	3 months
	PEN 5.00	USD 1.36	tomorrow	PEN 7.50	USD 2.03	3 months
	PEN 5.00	USD 1.36	tomorrow	PEN 10.00	USD 2.71	3 months
Choice set 2	PEN 5.00	USD 1.36	1 month	PEN 5.30	USD 1.44	4 months
	PEN 5.00	USD 1.36	1 month	PEN 6.00	USD 1.63	4 months
	PEN 5.00	USD 1.36	1 month	PEN 7.50	USD 2.03	4 months
	PEN 5.00	USD 1.36	1 month	PEN 10.00	USD 2.71	4 months
Choice set 3	PEN 5.00	USD 1.36	1 year	PEN 5.30	USD 1.44	1 year and 3 months
	PEN 5.00	USD 1.36	1 year	PEN 6.00	USD 1.63	1 year and 3 months
	PEN 5.00	USD 1.36	1 year	PEN 7.50	USD 2.03	1 year and 3 months
	PEN 5.00	USD 1.36	1 year	PEN 10.00	USD 2.71	1 year and 3 months

Source: Household Survey on the Impact of Floods in Lima (2023).

The order in which participants made their decisions was randomised on the level of the choice set. The sample is nearly evenly distributed across different sequences of choice sets (Table 4.A.1 in the Appendix). For the analysis of time preferences, I use the total number of patient choices made by the respondent, which is a simple count of how often the larger but later reward was chosen in all 12 choices.

Risk preferences — I followed the design created by Binswanger (1980), which has often been used in rural settings in developing countries (Bauer, Chytilova, and Morduch 2012; Chowdhury 2022). Respondents were presented with six investment options; each offering different levels of risk and return with equal probability (50%). The options ranged from a no-risk, guaranteed return to a high-risk option with the possibility of very high or very low returns. This setup allowed us to measure the participants' risk preferences by observing which options they chose, helping us understand their willingness to trade-off between security and potential gain. Table 4.2 shows the six investment options and the corresponding payoffs. For risk preferences, I used the gamble number picked as an outcome measure, a number from 1 to 6. Higher numbers are associated with a higher willingness to take risks. In Table 4.A.2 in the Appendix, I observe that almost 30% of households are extremely risk-averse.

Table 4.2 Risk Preferences: Payoffs offered to different investment options

Risk aversion levels	Investment option	Case A (Blue Marble)		Case B (Yellow Marble)	
Neutral-to-negative	1	PEN 0.00	USD 0.00	PEN 5.00	USD 1.36
Slight-to-neutral	2	PEN 0.40	USD 0.11	PEN 4.00	USD 1.08
Moderate	3	PEN 0.80	USD 0.22	PEN 3.50	USD 0.95
Intermediate	4	PEN 1.20	USD 0.33	PEN 3.00	USD 0.81
Severe	5	PEN 1.60	USD 0.43	PEN 2.50	USD 0.68
Extreme	6	PEN 2.00	USD 0.54	PEN 2.00	USD 0.54

Note: Household had to pick one out of six investment options.
Source: Household Survey on the Impact of Floods in Lima (2023).

As explained in Chapter 2, more recent and widely adopted methods—such as the Eckel-Grossman approach—offer comparable simplicity while addressing some of the limitations inherent in the Binswanger design, most notably by holding expected value constant across options. This feature allows for cleaner inference on risk preferences and more accurate parameter estimation. However, I chose to use the Binswanger approach due to its widespread application in similar empirical settings and its suitability for the study population.

Insurance experiment

A theoretical index-based flood insurance product was offered to each household in the sample. Participants listened to a brief outline of the Local Government's plans to introduce this insurance product, which provides compensation of PEN 10,000 (USD 2,711). The insurance pays out based on an index, in this case, rainfall, measured at a local weather station in the district of Lurigancho–Chosica. If the rainfall exceeds a pre-specified threshold known to cause devasting flooding (36 mm of accumulated rainfall), the index insurance pays policyholders the fixed amount. Households were free to express their willingness to pay (or not) for the hypothetical index-based flood insurance product. The index, based on the precipitation levels, relied on historical data from the Lurigancho–Chosica district's weather station[4].

Premium rates — I used a multiple monthly premium rates approach in which households reveal their willingness to pay (WTP) for index-based flood insurance to cover losses caused by a flood event. This hypothetical index-based insurance was priced at an actuarially fair price of PEN 227.27 (USD 61.61). First, they were randomly offered one out of five different prices to measure their willingness to purchase. Then, if a household did not choose to purchase the index-based insurance product, they received a price discount of 50 percent; otherwise, they faced a price increase of 50 percent. This method follows the double-bounded dichotomous choice contingent valuation (Hanemann et al., 1991) that improves efficiency over discrete choice models. I separate the probability of agreeing to pay insurance by first and follow-up question and different prices in Figure 4.A.1 in the Appendix.

Trust measures

To measure the level of trust in different types of insurance providers and the local government's city and flood management plans, I asked households the following questions: (1) In your opinion, which of the following institutions (private, NGO, or local government) do you trust the most to manage index insurance funds? (2) In your opinion, do you agree with the current local government's program for infrastructure development, education, and health? (3) Can the local government be trusted to effectively manage floods? and (4) Can the local government be trusted to manage flood insurance transparently? The first question is used to evaluate the households' expectation on a particular institution's responsibility to manage the funds and provide insurance to household in case it is implemented in the three urban communities of the district of Chosica, in Lima. The next question identifies the households' support of the current local government development program. The other two questions encompass a broader scope beyond index-based flood insurance. I then create three

[4] The National Service of Meteorology and Hydrology of Peru (SENAMHI) provided historical precipitation data from 2000 to 2022, which was used to identify periods of highest accumulated rainfall that caused devastating floods in Nicolás de Piérola, Sierra Limeña, and San Antonio de Pedregal.

dummy variables for trust towards the insurance institution and local government. I also create three dummy variables to represent trust towards the insurance institution and local government. Trust is coded as one if the respondent selects "completely agree" or "agree," and zero otherwise. In the Appendix, Figures 4.A.2 and 4.A.3 show the percentage of respondents who expressed trust in their type of insurer and local government, while Figure 4.A.4 presents trust levels in the local government and its flood plan across different measures of risk aversion. Overall, it is clear that trust is low across the three urban villages and at different levels of risk, particularly among households considered highly risk-averse.

Table 4.3 presents summary statistics for the sample of households, both as a whole and disaggregated by flood risk categories. Panel A displays the dependent variable, Panels B and C present the variables of interest, and Panels D to H provide the covariates.

4.4 Empirical Framework

4.4.1 Identification strategy

The study examines the effects of economic preferences on index-based insurance uptake. The exogenous variation was introduced by a lab-in-field experiment, which randomly offered households repeated discount rate payoffs to elicit time preferences and assigned monthly price bids to reveal the price they were willing to pay (i.e., insurance demand). Additionally, a gamble game was conducted, yielding either a high or low payoff with equal probability, to collect respondents' risk preferences.

The variables of interest are defined as follows to better understand their effects on insurance demand. First, time preference is measured by the number of times a household chooses Option 2, which offers a larger but delayed payoff. A higher number of patient choices indicates greater level of patience, making the household more likely to purchase insurance. Second, risk preference is assessed based on the respondent's selection from six investment options, each reflecting varying levels of risk aversion. Respondents who select the riskiest option (1: PEN 0.00 – PEN 5.00) are classified as risk-seeking, while those who choose the safest option (6: PEN 2.00 – PEN 2.00) are classified as extremely risk-averse. Extreme risk aversion is associated with a higher willingness to buy insurance. Third, price is defined as the natural logarithm of the insurance policy price, which includes the two-question price list. An increase in price leads to lower insurance demand.

The identification strategy exploits the exogeneity of the economic preferences variables, assuming they are exogenous after controlling for both the distance from the surveyed house to the arid riverbed and the urban community in which the house is located. The exogeneity of economic preferences has been supported by empirical studies, including those by Stigler and Becker (1977), Kundu (2022), Shin et al. (2022), Baillon et al. (2022), and Liu et al. (2020).

To determine whether the main variables are exogenous, I perform Ordinary Least Squares (OLS) regressions on several observables related to dwelling, household, and geographic characteristics, as well as experiences with floods (a total of 22 covariates). The independent variables are uncorrelated with all controls except for three at the 10% level of significance: number of storeys, female, and distance to weather station (see Table 4.A.3 in the Appendix). In this context, based on the assumption of selection on observables, I employ the Propensity Score Matching (PSM) estimator, which allows us to remove bias by conditioning on the propensity score $p(X_i)$ (Rosenbaum and Rubin, 1983) thereby obtaining fully exogenous variables of interest (Figure 4.A.5 in the Appendix shows the PSM graph, with observations lacking a counterfactual pair or off-support being dropped).

Table 4.3 Summary statistics

	(1) Flood-free areas (121 obs.)		(2) Flood-prone hillsides (140 obs.)		(3) Rainy season riverbeds (143 obs.)		(4) Total (404 obs.)			
	mean	sd	mean	sd	mean	sd	mean	sd	min	max
Panel A: Willingness-to-pay										
First question (1,0)	0.099	0.300	0.229	0.421	0.203	0.403	0.181	0.385	0	1
Follow-up question (1,0)	0.149	0.357	0.250	0.435	0.245	0.431	0.218	0.413	0	1
Panel B: Economic preferences										
Time Preference - Number of patient choices (in logs)	1.294	0.954	1.312	1.060	3.937	4.248	1.249	1.009	0	3
Risk Preference - Gamble number picked (number)	3.760	1.817	4.186	1.669	4.210	1.609	4.067	1.702	1	6
Risk Preference - Extreme risk aversion (1,0)	0.256	0.438	0.329	0.471	0.294	0.457	0.295	0.456	0	1
Price - First question (in logs)	5.000	1.384	4.425	1.588	4.590	1.480	4.655	1.506	2.398	6.217
Price - Follow-up question (in logs)	4.422	1.280	3.938	1.462	4.122	1.378	4.148	1.390	1.792	6.909
Panel C: Trust										
Trust local government (1,0)	0.455	0.500	0.264	0.443	0.441	0.498	0.384	0.487	0	1
Trust local government flood plan (1,0)	0.430	0.497	0.307	0.463	0.399	0.491	0.376	0.485	0	1
Panel D: Dwelling characteristics										
Electricity (1,0)	0.975	0.156	0.986	0.119	0.986	0.118	0.983	0.131	0	1
Water (1,0)	0.174	0.380	0.143	0.351	0.175	0.381	0.163	0.370	0	1
Sewage (1,0)	0.950	0.218	0.921	0.270	0.930	0.256	0.933	0.250	0	1
Toilet (1,0)	0.992	0.091	0.993	0.085	0.993	0.084	0.993	0.086	0	1
Floor (1,0)	0.967	0.180	0.979	0.145	0.986	0.118	0.978	0.148	0	1
Number of storeys	1.802	0.843	1.629	0.817	1.741	0.776	1.720	0.812	1	4
Have land title (1,0)	0.397	0.491	0.350	0.479	0.287	0.454	0.342	0.475	0	1

Panel E: Household characteristics										
Household size (in logs)	1.582	0.370	1.602	0.370	1.599	0.380	1.595	0.372	0.693	2.773
Age (in logs)	3.871	0.307	3.872	0.328	3.939	0.316	3.895	0.319	2.996	4.394
Education (in years)	14.008	3.902	13.671	4.254	14.203	4.440	13.960	4.215	0	21
Self-/employed worker last week (1,0)	0.273	0.447	0.200	0.401	0.210	0.409	0.225	0.418	0	1
Female (1,0)	0.694	0.463	0.614	0.489	0.727	0.447	0.678	0.468	0	1
Married (1,0)	0.314	0.466	0.264	0.443	0.322	0.469	0.300	0.459	0	1
Monthly expenditure (in logs)	6.959	1.310	6.734	1.100	6.814	1.288	6.830	1.233	0	8.854
Public health insurance (1,0)	0.826	0.380	0.764	0.426	0.811	0.393	0.800	0.401	0	1
Panel F: Flood experience, costs and shocks										
Flood experience in 2015-2023 (1,0)	0.868	0.340	0.900	0.301	0.727	0.447	0.829	0.377	0	1
Flood expenditure in 2017 (in logs)	0.000	0.000	0.784	1.987	0.470	1.593	0.438	1.535	0	7.445
Floods in last 10 years (number)	1.711	1.136	1.486	0.917	1.336	1.113	1.500	1.065	0	3
Panel G: Geographic characteristics										
Distance to watercourse (in logs)	5.096	0.611	4.943	0.695	3.843	0.566	4.600	0.842	2.139	6.091
Distance to weather station (in km)	2.087	0.612	2.038	0.617	2.130	0.687	2.085	0.641	1.283	3.320
Distance to geodynamic barrier (in logs)	7.018	0.468	6.988	0.446	7.152	0.319	7.055	0.418	5.601	7.683
Elevation (in logs)	6.837	0.144	6.853	0.141	6.840	0.078	6.844	0.123	6.208	7.004
Panel H: Household mitigation										
At least one mitigation measure (1,0)	0.554	0.499	0.686	0.466	0.664	0.474	0.639	0.481	0	1

Source: Household Survey on the Impact of Floods in Lima (2023).

Table 4.4 Exogeneity tests: OLS regressions after PSM

Sample: **Dependent variable**	**Number of patient choices (in logs)**		**Extreme risk aversion (1,0)**		**Premium (in logs)**	
Electricity (1,0)	-0.0089	(0.0065)	0.0078	(0.0147)	-0.0050	(0.0044)
Water (1,0)	-0.0049	(0.0181)	-0.0142	(0.0408)	-0.0096	(0.0123)
Sewage (1,0)	-0.0105	(0.0115)	0.0362	(0.0258)	0.0099	(0.0078)
Toilet (1,0)	-0.0030	(0.0043)	0.0129	(0.0096)	-0.0006	(0.0029)
Floor (1,0)	-0.0004	(0.0074)	-0.0228	(0.0166)	-0.0063	(0.0050)
Storey (number)	0.0046	(0.0387)	-0.1430	(0.0867)	-0.0363	(0.0261)
Have land title (1,0)	0.0008	(0.0228)	-0.0168	(0.0513)	-0.0124	(0.0154)
Female (1,0)	-0.0039	(0.0233)	0.0840	(0.0523)	0.0221	(0.0157)
Age (in logs)	-0.0084	(0.0157)	0.0037	(0.0353)	-0.0138	(0.0106)
Married (1,0)	0.0070	(0.0230)	0.0067	(0.0517)	-0.0236	(0.0155)
Education (years)	-0.2006	(0.2094)	-0.1719	(0.4714)	-0.0728	(0.1417)
Household size (in logs)	0.0153	(0.0186)	0.0063	(0.0419)	-0.0015	(0.0126)
Self-/employed last week (1,0)	0.0139	(0.0210)	-0.0007	(0.0473)	0.0075	(0.0142)
Household expenditure (in logs)	-0.0578	(0.0617)	-0.1827	(0.1388)	0.0482	(0.0417)
Public health insurance (1,0)	-0.0062	(0.0201)	-0.0061	(0.0453)	0.0213	(0.0136)
Flood experience 2015-2023 (1,0)	-0.0070	(0.0186)	0.0293	(0.0419)	0.0152	(0.0126)
Flood expenditure 2017 (in logs)	-0.0017	(0.0531)	-0.0602	(0.1194)	0.0248	(0.0359)
Floods in last 10 years (number)	0.0620	(0.0762)	0.0406	(0.1717)	0.0497	(0.0515)
Household mitigation (1,0)	0.0077	(0.0240)	-0.0407	(0.0541)	0.0039	(0.0163)
Distance to weather station (km)	-0.0142	(0.0087)	0.0176	(0.0197)	-0.0059	(0.0059)
Distance to barrier (in logs)	-0.0256	(0.0198)	0.0036	(0.0447)	-0.0074	(0.0134)
Elevation (in logs)	0.0000	(0.0059)	0.0105	(0.0133)	0.0014	(0.0040)
Distance to watercourse (in logs)	Yes		Yes		Yes	
Urban community FE	Yes		Yes		Yes	
Number of observations (N)	400		400		400	
Test of joint significance	F-stat: 0.63 (p-value: 0.9004)		F-stat: 0.78 (p-value: 0.7524)		F-stat: 0.82 (p-value: 0.7064)	

Note: OLS regression estimates for continuous and dummy variables. Each result represents a regression of the dependent variable on an explanatory variable (which are the study variables). Each result displays the coefficients related to each explanatory variable of interest and the standard errors in parenthesis. Each regression includes distance to riverbed and urban community fixed effect. Asterisk (*), double asterisk (**), and triple asterisk (***) denote variables significant at 10 percent, 5 percent, and 1 percent, respectively.
Source: Household Survey on the Impact of Floods in Lima (2023).

The matching process resulted in the discarding of only four observations, leaving a final sample of 400 observations. As a result of the matching, there are no longer any covariates significantly associated with any of the variables of interest (see Table 4.4). I then rely on the dataset based on the new weights constructed from the sample weights and matching results, as the covariates are no longer significantly associated with the main variables, achieving exogeneity.

Table 4.5 Correlation of economic preferences

Variable of interest	(1) Number of patient choices (in logs)	(2) Extreme aversion (1,0)	(3) Premium rates (in logs)
Number of patient choices (in logs)	1.000		
Extreme risk averse (1,0)	-0.056	1.000	
Premium (in logs)	-0.048	-0.014	1.000

Source: Household Survey on the Impact of Floods in Lima (2023).

In addition, I examine the correlation between the main variables. Table 4.5 presents correlations of economic preferences. I see that they are insignificantly negatively correlated between time and risk preferences, and premium rates.

Table 4.6 Balance tests between flood-free and flood-prone urban communities after PSM

	(1) Flood-free group	(2) Flood-prone group	(3) Single difference	(4) N	(5) p-value	
A. Variables of interest						
Number of patient choices (in logs)	1.294	1.217	0.077	400	0.485	
Extreme aversion (1,0)	0.256	0.312	-0.056	400	0.264	
Premium - First question (in logs)	5.000	4.511	0.489	400	0.003	***
B. Covariates						
Electricity (1,0)	0.975	0.986	-0.010	400	0.465	
Water (1,0)	0.174	0.161	0.012	400	0.762	
Sewage (1,0)	0.950	0.939	0.011	400	0.655	
Toilet (1,0)	0.992	0.993	-0.001	400	0.907	
Floor (1,0)	0.967	0.982	-0.015	400	0.350	
Storey (number)	1.802	1.688	0.113	400	0.199	
Have land title (1,0)	0.397	0.315	0.081	400	0.116	
Female (1,0)	0.694	0.677	0.017	400	0.741	
Age (in logs)	3.871	3.910	-0.039	400	0.259	
Married (1,0)	0.314	0.297	0.017	400	0.741	
Education (years)	14.008	13.907	0.101	400	0.826	
Household size (in logs)	1.582	1.595	-0.013	400	0.741	
Self-/employed last week (1,0)	0.273	0.208	0.065	400	0.156	
Household expenditure (in logs)	6.959	6.763	0.196	400	0.145	
Public health insurance (1,0)	0.826	0.785	0.042	400	0.344	
Distance to weather station (km)	2.087	2.084	0.003	400	0.961	
Distance to barrier (in logs)	7.018	7.076	-0.058	400	0.200	
Elevation (in logs)	6.837	6.848	-0.011	400	0.398	

Note: This table presents the summary statistics (standard errors in brackets) for the sample used in this study, along with the p-values for differences in means between the randomly assigned flood-free or flood-prone groups.
Source: Household Survey on the Impact of Floods in Lima (2023).

Using a two-sample t-test, I conduct balance tests to determine whether the groups in the sample are balanced[5]. Table 4.6 reports the balance tests on the assignment to flood-prone and flood-free areas among the surveyed households in the district of Chosica. There is no significant difference in outcomes (a total of 3) and covariates (a total of 18) between households located in flood-risk and non-flood-risk areas, except for premium rates in the first question, which is statistically significant at the 1% level. Although insurance prices are unbalanced, I confirm that the randomisation was successful during data collection since all covariates are well-balanced.

Finally, the identification strategy is independent of observed controls in estimating the effects of economic preferences on demand. In the analysis below, the results remain robust regardless of whether household-level covariates are included in the estimations. This is expected, given that the balance tests indicate no significant differences across household characteristics between treatment (i.e., flood-prone) and control (i.e., flood-free) groups.

4.4.2 Empirical equation

The demand model is

$$Q_{ij} = \beta_0 + \boldsymbol{\beta_1}\boldsymbol{\sigma_i} + \boldsymbol{\beta_2}\boldsymbol{\lambda_i} + \boldsymbol{\beta_3}\boldsymbol{\alpha_i} + X_i'\boldsymbol{\beta_X} + \gamma G_{ij} + \lambda F_i + \theta_j + \phi_r + \varepsilon_{i,j} \qquad (1)$$

In equation (1), Q_{ij} is the willingness to pay made by household *i* in the urban community *j* (a binary variable = 1 if the *i*th household in the urban community *j* is willing to purchase the product, and = 0 otherwise). The economic preference parameters include $\boldsymbol{\sigma_i}$ (time preference), $\boldsymbol{\lambda_i}$ (extreme risk aversion) and $\boldsymbol{\alpha_i}$ (first and follow-up question premium rates). Let $\boldsymbol{X_i}$ be the vector of dwelling and household characteristics; G_i indicates the vector of geographical characteristics in the urban community; and I control for distance from the surveyed house to the watercourse (F_i), urban community (θ_j), and flood exposure (ϕ_r); and $\varepsilon_{i,j}$ is the random error.

I have two observations for each of the 400 surveyed household (i.e., one for each premium level), giving us a total sample size of 800 observations. Equation (1) is estimated by ordinary least squares (OLS), as are all other regressions. I bootstrap 1,000 replications to all OLS

[5] I present balance test results estimated on the set of outcomes and covariates using the full sample in Table 4.A.4 in the Appendix. The results are very similar to the balanced estimates.

regressions to approximate standard errors, confidence intervals, and p-values for test statistics, based on the sample data[6].

4.5 Results

4.5.1 Main results

This section presents the findings from the lab-in-field experiment, which investigates the effects of time and risk preferences, elicited through discount rate and risky games, as well as random variation in premium prices on index-based insurance demand.

In column (1) of Table 4.7, I observe that demand for index-based flood insurance is positively and statistically significant at the 5 percent level for the number of patient choices (indicating weaker time preferences). Here the estimate indicates that a standard deviation increase in the time preference measure is associated with a 2.5-pp increase in the probability of insurance, on average[7]. Additionally, premium rates are negatively and statistically significant at the 1 percent level. A 10% decrease in insurance price seems to lead to a 0.9 percentage point increase in take-up[8], which, given the low levels of effective demand (11.4%), corresponds to an 8.2% increase in take-up[9]. I also estimate the relationship between risk aversion and demand for index-based insurance. Although I do not find a significant effect of risk aversion on demand, I infer that as household risk aversion increases, the probability of being insured decreases.

Similarly, column (2) shows that time preference is statistically significant at the 5 percent level, while premium remains significant at the 1 percent level. In this specification, I analyse the effect of extreme risk aversion rather than levels of risk aversion. An extremely risk-averse household exhibits a negative and statistically significant relationship with insurance demand at the 1 percent level. I estimate that a standard deviation increase in extreme risk-aversion is associated with a 3.6 percentage point decrease in the probability of purchasing index-based insurance.

---

6 The sample size, while sufficient for the analysis, may not be large enough for traditional asymptotic approximations to be fully accurate. Bootstrapping helps to mitigate this issue by generating an empirical distribution of the estimator from the data.

7 This is calculated by the product of the OLS marginal effect for the "number of patient choices (in logs)" variable in column (1) of Table 7, 0.0248, and the standard deviation of that variable (SD 1.009).

8 This follows from the product of the OLS marginal effect for the ''Premium (in logs)'' variable in column (1) of Table 7, -0.0978, and the natural logarithm transformation of a 10% increase in said variable, in (1.1).

9 This is calculated by the percentage increase in take-up (0.94 pp) relative to the effective demand of 11.4%.

Thus far, I have considered the impact of the variables of interest in isolation. However, I am likely to observe different price elasticity for households with a strong time preference—that is, a tendency for immediate payoffs—compared to households with weaker time preferences, who focus on the future and delay rewards (such as insurance compensation). I test this assumption in column (3) by interacting the number of patient choices with the premium. The results confirm this assumption: the price sensitivity of demand increases with a household's level of patience (see Figure 4.A.6 in the Appendix). In other words, a more patient household is sensitive to increases in premium, resulting in a decrease in insurance take-up as premiums rise. Moreover, subsidies are more effective in stimulating demand when demand shows greater price sensitivity.

In column (4), I also examine whether price elasticity varies across levels of risk aversion. The interaction term suggests that extremely risk-averse households are less price-sensitive and, as in column (2), less likely to purchase index-based insurance. This implies that even if extreme risk-averse households receive subsidies, they are still unlikely to buy the insurance product. In Figure 4.A.7 in the Appendix, I plot the average marginal effect of the logarithm of price on the probability of uptake across households with different levels of risk aversion. I observe that risk-seeking (neutral-to-negative) households have price sensitivity 1.62 times higher than that of extremely risk-averse households. Finally, in column (5), the interaction term between extreme risk aversion and time preference is statistically insignificant, confirming that both variables are fully exogenous.

Additional robustness tests are undertaken to check on the exogeneity of the estimated coefficients. I perform an Oster test, as proposed by Oster (2019), to evaluate the stability of the effects of economic preferences on insurance demand in the presence of unobservable factors. I set the coefficient of proportionality $\boldsymbol{\delta}$ to 1, which assumes that "the observables are at least as important as the unobservables" (Oster, 2019: 195–196) in determining the outcome variables. Additionally, I increase the value of R-maximum to 30% above the R-squared value from the main model regression. The results of the Oster test indicate that the coefficients for the number of patient choices, extreme risk aversion, and premium are relatively consistent with those in the main model. These findings suggest that the results are unlikely to be biased by unobservables (see Panel B in Table 4.7).

To confirm the validity of the main findings, I specifically conduct two alternative specifications—a Probit model and an OLS regression without covariates—to determine whether the primary results are consistent in magnitude and sign with those obtained from the OLS specification. Table 4.8 reports that Probit coefficients are higher than OLS estimates but maintain the same signs. OLS and Probit estimates share the same sign because they capture the same direction of the relationship between variables; however, they differ in magnitude due to differences in model assumptions and the nature of the probability distribution used in Probit. On the other hand, the OLS estimates excluding covariates are relatively identical in both magnitude and sign to the main OLS estimates.

Table 4.7 Index-based flood insurance demand among surveyed households

Model: OLS	Dependent variable: Insurance take-up									
	(1)		(2)		(3)		(4)		(5)	
PANEL A										
Number of patient choices (in logs)	0.025	**	0.024	**	0.135	***	0.023	*	0.036	**
	(0.012)		(0.012)		(0.050)		(0.012)		(0.016)	
Extreme aversion (number)	-0.008									
	(0.008)									
Extreme risk aversion (1,0)			-0.078	***	-0.072	**	-0.227	**	-0.038	
			(0.028)		(0.029)		(0.105)		(0.042)	
Premium (in logs)	-0.098	***	-0.098	***	-0.066	***	-0.108	***	-0.097	***
	(0.010)		(0.011)		(0.016)		(0.013)		(0.011)	
Number of patient choices (in logs) x Premium (in logs)					-0.025	**				
					(0.010)					
Extreme risk aversion (1,0) x Premium (in logs)							0.034	*		
							(0.020)			
Number of patient choices (in logs) x Extreme risk aversion (1,0)									-0.034	
									(0.026)	
PANEL B										
Bias-adjusted estimates, with Oster δ=1 and $Rmax$=1.3$\tilde{R}$										
- Number of patient choices (in logs)	0.024		0.023		-		-		-	
- Extreme aversion (number)	-0.008		-		-		-		-	
- Extreme risk aversion (1,0)	-		-0.078		-		-		-	
- Premium (in logs)	-0.102		-0.102		-		-		-	
Distance to watercourse (in logs)	Yes		Yes		Yes		Yes		Yes	
Urban community FE	Yes		Yes		Yes		Yes		Yes	
Flood exposure FE	Yes		Yes		Yes		Yes		Yes	
Covariates	Yes		Yes		Yes		Yes		Yes	
Number of observations	800		800		800		800		800	

Note: This table presents OLS regression estimates of household-level insurance uptake on economic preferences, based on double-bounded dichotomous choice questions. All specifications include distance to watercourse, covariates, urban community and flood exposure fixed effects. Standard errors are bootstrapped (1,000 replications). Asterisk (*), double asterisk (**), and triple asterisk (***) denote variables significant at 10 percent, 5 percent, and 1 percent, respectively. Source: Household Survey on the Impact of Floods in Lima (2023).

Table 4.8 Probit specification and OLS model without covariates

	Dependent variable: Insurance take-up							
Model:	Probit				OLS			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Number of patient choices (in logs)	0.115 *	0.462 **	0.114 *	0.140 *	0.020 *	0.098 *	0.019	0.024
	(0.065)	(0.203)	(0.063)	(0.077)	(0.012)	(0.050)	(0.013)	(0.017)
Extreme risk aversion (1,0)	-0.390 ***	-0.349 **	-0.476	-0.288	-0.076 ***	-0.073 **	-0.208 *	-0.061
	(0.142)	(0.155)	(0.428)	(0.234)	(0.028)	(0.029)	(0.106)	(0.042)
Premium (in logs)	-0.416 ***	-0.306 ***	-0.422 ***	-0.413 ***	-0.091 ***	-0.069 ***	-0.101 ***	-0.091 ***
	(0.048)	(0.077)	(0.053)	(0.051)	(0.010)	(0.016)	(0.012)	(0.010)
Number of patient choices (in logs) x Premium (in logs)		-0.088 *				-0.017 *		
		(0.049)				(0.010)		
Extreme risk aversion (1,0) x Premium (in logs)			0.023				0.030	
			(0.103)				(0.020)	
Number of patient choices (in logs) x Extreme risk-averse (1,0)				-0.085				-0.013
				(0.141)				(0.026)
Distance to watercourse (in logs)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Urban community FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Flood exposure FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Covariates	Yes	Yes	Yes	Yes	No	No	Yes	Yes
Number of observations	800	800	800	800	800	800	800	800

Note: This table presents Probit and OLS regression estimates of household-level insurance uptake on economic preferences, based on double-bounded dichotomous choice questions. Marginal effects are reported in columns (1) to (4). All specifications include distance to watercourse, covariates, urban community and flood exposure fixed effects. Standard errors are bootstrapped (1,000 replications). Asterisk (*), double asterisk (**), and triple asterisk (***) denote variables significant at 10 percent, 5 percent, and 1 percent, respectively.
Source: Household Survey on the Impact of Floods in Lima (2023).

4.5.2 Extended analysis

In this section, I extend the investigation by taking a closer look at the estimates of time and risk preferences and their interaction terms with the premium.

Time preference

Figure 4.3 presents the estimated probability of purchasing index-based insurance (with 95% confidence intervals) as a function of the number of patient choices (in logs). The conditional mean function is estimated using a bivariate local linear regression (Fan and Gijbels, 1996), which makes no assumptions about the nature of the underlying relationship[10]. I observe that households with a higher number of patient choices (indicating weaker time preference) have the highest probability of purchasing the insurance product. While the 95% confidence intervals overlap, the point estimates suggest that the probability of purchasing index-based insurance increases as the number of patient choices rises (indicating weaker time preference) over much of its range, before declining at around 7 patient choices.

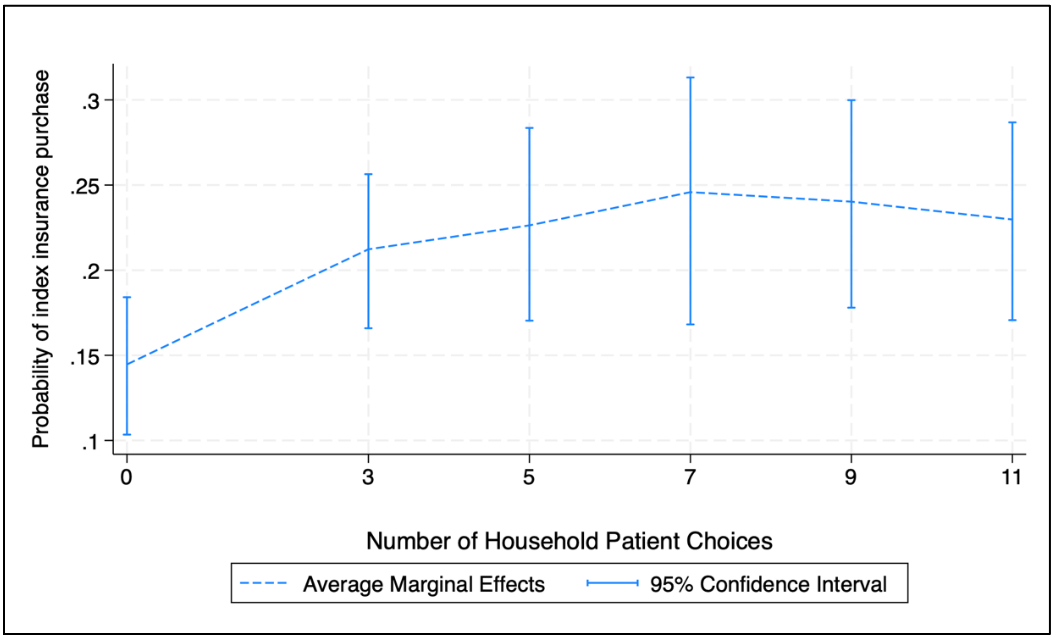


Figure 4.3 Probability of index insurance as a function of the non-parametric time preference

Note: This figure shows the probability of purchasing index-based insurance as a function of nonparametric time preference measure (number of patient choices). The graph is obtained from a bivariate local linear regression of an insurance indicator (willingness to pay for hypothetical index-based insurance) on the respective time preference measure. The Epanechnikov kernel function is used and the bandwidth is selected by the plugin estimator of the asymptotically optimal constant bandwidth. Whiskers show 95% confidence intervals (CI).
Source: Household Survey on the Impact of Floods in Lima (2023).

[10] The conditional mean functions are robust to estimation by local constant regression (Nadaraya, 1965; Watson, 1964) rather than local linear regression.

Now I construct two dummy variables that represent low and high patience choices. This distinction allows us to separately estimate the relationship between the extreme cases of time preference and demand. Again, I see the estimate of low patient choices is negative and statistically significant at 5% in column (1) of Table 4.9. This implies impatient households are unlikely to purchase index-based insurance. In column (3), the interaction coefficient between low patient choices and expenditure is negative and statistically significant at 10% suggesting that even as spending capacity increases, low-patient households may still be less likely to purchase insurance, possibly due to the household's preference for immediate benefits over future ones. Conversely, I infer that for households with low patience and low expenditure, the willingness to pay for insurance increases less with additional expenditure compared to more patient households.

Table 4.9 Effects of levels of patience and expenditure on insurance demand

Model: OLS	Dependent variable: Insurance take-up							
	(1)		(2)		(3)		(4)	
Number of patient choices:								
- Low patient choices (1,0)	-0.052	**			0.025			
	(0.028)				(0.049)			
- High patient choices (1,0)			0.047				-0.018	
			(0.032)				(0.053)	
Extreme risk-averse (1,0)	-0.082	***	-0.081	***	-0.079	***	-0.080	***
	(0.029)		(0.028)		(0.028)		(0.029)	
Premium (in logs)	-0.099	***	-0.099	***	-0.100	***	-0.099	***
	(0.010)		(0.010)		(0.010)		(0.010)	
Low patient choices (1,0) x Expenditure					0.000	*		
					(0.000)			
High patient choices (1,0) x Expenditure							0.0000	
							(0.000)	
Distance to watercourse (in logs)	Yes		Yes		Yes		Yes	
Urban community FE	Yes		Yes		Yes		Yes	
Flood exposure FE	Yes		Yes		Yes		Yes	
Covariates	Yes		Yes		Yes		Yes	
Number of observations	800		800		800		800	

Note: This table presents OLS regression estimates of household-level insurance uptake based on economic preferences, using double-bounded dichotomous choice questions. The number of patient choices comes from the time preference game, where "Low patient choices" is a dummy variable that takes the value of 1 if the choice belongs to one of the first five options (0 to 4), and 0 otherwise. "High patient choices" is a dummy variable that takes the value of 1 if the choice belongs to one of the last five options (8 to 12) and 0 otherwise. All specifications include distance to watercourse, covariates, urban community and flood exposure fixed effects. Standard errors are bootstrapped (1,000 replications). Asterisk (*), double asterisk (**), and triple asterisk (***) denote variables significant at 10 percent, 5 percent, and 1 percent, respectively.
Source: Household Survey on the Impact of Floods in Lima (2023).

I also asked households about the key factors needed for index-based insurance to be effective if implemented in the district of Lurigancho–Chosica. In column (5) of Table 4.10, the coefficient for low-patience households is positive and statistically significant at the 10% level. It suggests that households with a strong time preference (low patience) are sensitive to the waiting period between compensation approval and fund disbursement if a flood occurs. This finding becomes crucial for the case since many households in the sample face liquidity constraints, with 70% categorised as poor and vulnerable, lacking available cash on hand.

In other words, people with a high time preference, who prefer immediate rewards, are less likely to buy insurance. This aligns with economic theory: if newly insured individuals must wait to make a claim and cannot borrow at low rates to cover the insurance premium, they may avoid purchasing insurance altogether.

Table 4.10 Factors to determine the success of implementation of index-based flood insurance in the district of Lurigancho–Chosica

	Dependent variable: If insurance is implemented, what factors are the most important for it to be successful?									
Model: OLS	(1) Insurance provider		(2) Compensation amount		(3) Insurance price amount		(4) Compensation approval process		(5) Compensation disbursement time	
Number of patient choices:										
- Low patient choices (1,0)	-0.079 **		0.012		0.007		0.033		0.040 *	
	(0.032)		(0.026)		(0.035)		(0.022)		(0.023)	
- High patient choices (1,0)		0.031		-0.021		0.037		-0.009		-0.047 **
		(0.035)		(0.029)		(0.038)		(0.025)		(0.024)
Extreme risk-averse (1,0)	-0.036	-0.035	-0.039	-0.040	0.028	0.028	0.033	0.032	0.024	0.023
	(0.033)	(0.034)	(0.027)	(0.028)	(0.037)	(0.038)	(0.022)	(0.022)	(0.024)	(0.024)
Premium (in logs)	-0.007	-0.008	-0.009	-0.009	0.049 ***	0.050 ***	-0.023 ***	-0.022 ***	-0.008	-0.008
	(0.010)	(0.011)	(0.008)	(0.008)	(0.011)	(0.011)	(0.007)	(0.007)	(0.007)	(0.008)
Distance to watercourse (in logs)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Urban community FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Flood exposure FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Covariates	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Number of observations	800	800	800	800	800	800	800	800	800	800

Note: This table presents OLS regression estimates of household-level insurance uptake based on economic preferences, using double-bounded dichotomous choice questions. The number of patient choices comes from the time preference game, where "Low patient choices" is a dummy variable that takes the value of 1 if the choice belongs to one of the first five options (0 to 4), and 0 otherwise. "High patient choices" is a dummy variable that takes the value of 1 if the choice belongs to one of the last five options (8 to 12) and 0 otherwise. All specifications include distance to watercourse, covariates, urban community and flood exposure fixed effects. Standard errors are bootstrapped (1,000 replications). Asterisk (*), double asterisk (**), and triple asterisk (***) denote variables significant at 10 percent, 5 percent, and 1 percent, respectively.
Source: Household Survey on the Impact of Floods in Lima (2023).

Risk preferences

Figure 4.4 presents the estimated probability of purchasing index-based insurance, along with 95% confidence intervals, as a function of varying levels of risk aversion. The conditional mean function is estimated using a bivariate local linear regression (Fan and Gijbels, 1996), which allows for flexibility in capturing the underlying relationship without assuming a specific functional form.

I observe a hump-shaped demand pattern: households with moderate to intermediate levels of risk aversion exhibit the highest WTP for index-based flood insurance. Beyond this peak, the demand curve slopes downward, with the lowest WTP observed among households that are extremely risk-averse. Notably, the probability of being insured is higher for households who are neutral-to-negative (i.e., risk-seeking) than for those who are extremely risk-averse. A distinct "kink" appears in the demand curve for households with intermediate levels of risk aversion.

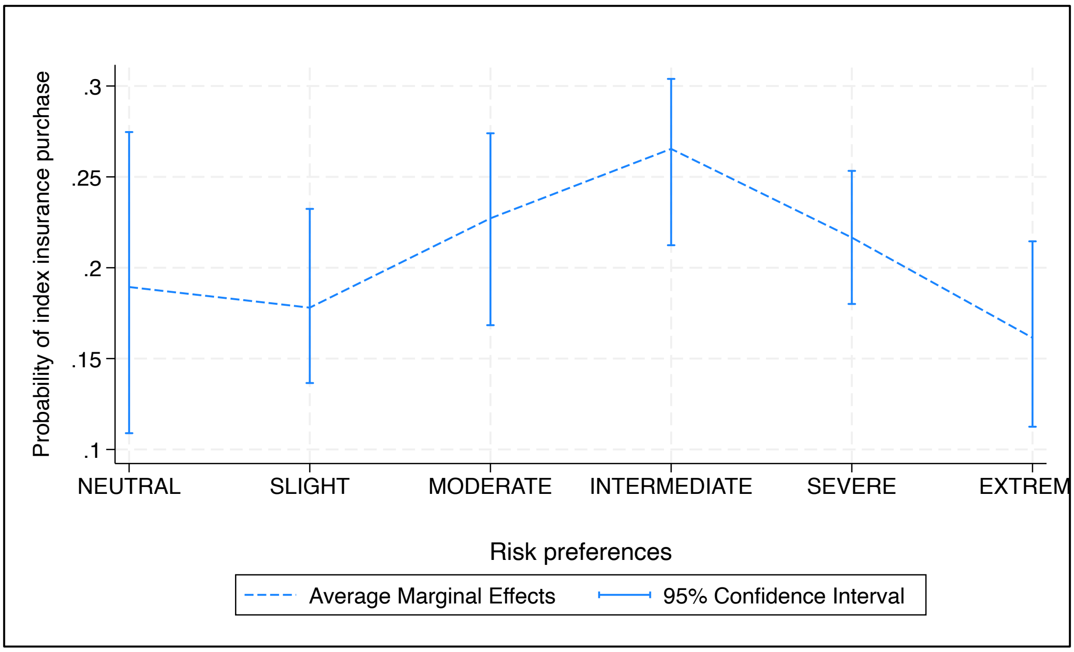


Figure 4.4 Probability of index-based insurance as a function of the non-parametric risk aversion levels

Note: This figure shows the probability of purchasing index-based insurance as a function of nonparametric risk preference measure (risk aversion levels). The graph is obtained from a bivariate local linear regression of an insurance indicator (willingness to pay for a hypothetical index insurance) on the respective time preference measure. The Epanechnikov kernel function is used and the bandwidth is selected by the plugin estimator of the asymptotically optimal constant bandwidth. Whiskers show 95% confidence intervals (CI).
Source: Household Survey on the Impact of Floods in Lima (2023).

Lastly, I analyse the changes in the effects of levels of risk aversion (particularly, neutral-to-negative and extreme risk aversion) on demand concerning changes in premium levels. Figure 4.5 shows that the impact of risk-seeking households on demand is negative and remains below zero as the premium increases. In contrast, the effects of least risk-averse households on demand are positive; however, beyond a premium of 50 PEN (USD 13.55), the effects become negative and remain below zero as the premium increases.

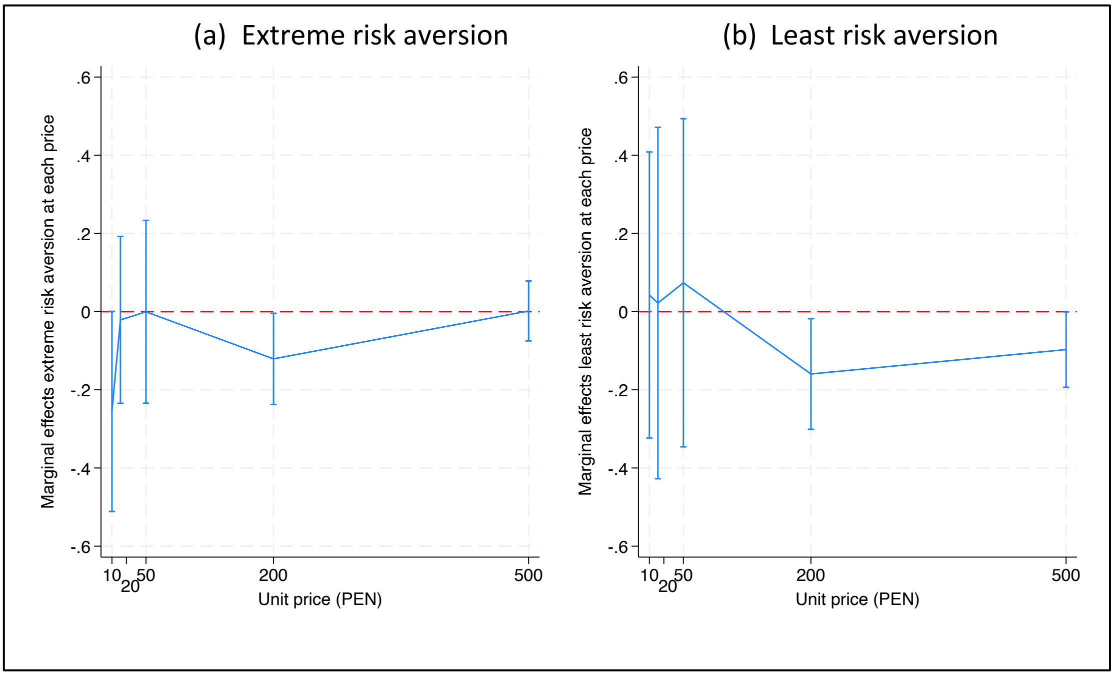


Figure 4.5 Marginal effects of extreme and least risk aversion on demand at different prices

Note: These figures illustrate the effects of extreme risk aversion and risk-seeking behaviour on demand in response to changes in premium levels, using coefficients from estimating equation (1). Figure 4.5 (a) shows the negative effect of extreme risk aversion on insurance demand as the premium households face rises from 10 PEN/unit. Figure 4.5 (b) demonstrates that the effect of least risk aversion on insurance demand is positive as the premium increases but begins to decrease when the premium exceeds 50 PEN/unit.
Source: Household Survey on the Impact of Floods in Lima (2023).

4.5.3 Trust as a mechanism

To this end, I have examined the effects of the economic preferences on the uptake for index-based insurance. Now, this section will elaborate further on the possible mechanisms underlying the main findings. I replicate the main specification using trust outcomes as a dependent variable. I check whether the level of trust explains the low demand and has any effects on insurance uptake.

For this reason, I use the following OLS specification:

$$Q_{ij} = \beta_0 + \boldsymbol{\beta_1 \sigma_i} + \boldsymbol{\beta_2 \lambda_i} + \boldsymbol{\beta_3 \alpha_i} + \boldsymbol{\beta_4 \eta_i} + X_i' \boldsymbol{\beta_X} + \gamma G_{ij} + \lambda F_i + \phi_r + \theta_j + \varepsilon_{i,j} \quad (2)$$

where I introduce $\boldsymbol{\eta}_i$ which represents the level of trust of ith household, along with the economic preference variables, the vector of geographical and household characteristics, distance from the surveyed household to the watercourse, urban community and flood exposure fixed effects, and random error.

Table 4.11 reports the results from estimating economic preferences and trust for three different institutions that could potentially manage and offer index-based insurance to households in the district of Chosica. Each model controls for the house's distance to the watercourse, urban community fixed effects, and flood risk exposure. In column (1), I find that greater patience (as measured by the number of patient choices made by the household) has a negative effect on trust in the government as an index-based insurance provider, significant at the 5% level. Similarly, extremely risk-averse households show a preference against the local government as an insurance provider, although this effect is statistically insignificant.

Additionally, when premium rates are higher, support for government-provided insurance increases, likely because government involvement (from households' point of view) is seen as a safer option with the capacity to offer subsidies or regulate premiums more effectively, thereby reducing financial burdens for the insured.

In column (2), I observe contrasting results in the estimates for levels of patience and risk aversion. It shows that the more patient trust more in private institutions (at 5% significance level) whereas the more risk averse trust less in insurance companies (at 1% significance level). Column (3) shows that trust in NGOs is negatively associated with patience at the 10% significance level, while extremely risk-averse households tend to trust NGOs more at the 5% level.

Although the two economic preference estimates are opposed in sign, it is shown that the magnitude and significance level of the risk aversion are higher than that of the estimated time preference. This places less trust in the efficiency and reliability of private insurance companies and more trust in NGOs to manage and provide the index-based insurance.

This may be due to the perception that private companies offer more tailored or financially stable products for long-term benefits. This suggests that more patient households, with a preference for delayed gratification, seek private-sector solutions they deem more sustainable. In contrast, extreme risk aversion significantly shapes preferences for insurance providers.

Highly risk-averse households tend to avoid private companies, likely perceiving them as too profit-oriented or volatile in crisis situations. Instead, these households show a strong preference for NGO-provided insurance, which they may associate with humanitarian or mission-driven approaches.

Columns (4) to (6) in Table 4.11 show the reasons why households are unlikely to purchase index-based insurance. In column (4), I observe that extremely risk-averse households believe they do not need index-based insurance. They may feel they do not need insurance, potentially viewing it as a cost today if they may not receive the indemnity in the future. Moreover, as premium rates increase the need for the insurance product decreases.

In columns (5) and (6), I find that more patient households tend to be unwilling to purchase index-based insurance if they face liquidity constraints, as financial limitations influence their decisions regarding insurance products that promise a future payout under uncertainty. Furthermore, such households often express a preference for private institutions to handle the insurance products, reflecting their trust in formal financial intermediaries and a desire for greater control over their assets. In contrast, the extremely risk-averse household estimate correlates negatively with a lack of trust in the insurance system as indicated in the previous table.

Table 4.11 Trust in difference institutions to manage and offer index-based insurance and reasons why households are unwilling to purchase insurance

Dependent variable:	Would you buy if insurance is provided by:			Would you not buy insurance because there is:		
Model: OLS	(1) Local Government	(2) Insurance Company	(3) Int'l Org. / NGO	(4) No need of insurance	(5) Liquidity constrained	(6) No trust in insurance system
Number of patient choices (in logs)	-0.030 **	0.034 **	-0.027 *	0.000	0.037 **	-0.044 ***
	(0.015)	(0.014)	(0.016)	(0.013)	(0.016)	(0.011)
Extreme risk aversion (1,0)	-0.014	-0.097 ***	0.093 **	0.061 *	-0.045	0.040 *
	(0.033)	(0.032)	(0.039)	(0.032)	(0.034)	(0.024)
Premium (in logs)	0.019 *	-0.009	-0.013	0.018 *	0.124 ***	-0.001
	(0.010)	(0.010)	(0.011)	(0.010)	(0.010)	(0.007)
Distance to watercourse (in logs)	Yes	Yes	Yes	Yes	Yes	Yes
Urban community FE	Yes	Yes	Yes	Yes	Yes	Yes
Flood exposure FE	Yes	Yes	Yes	Yes	Yes	Yes
Covariates	Yes	Yes	Yes	Yes	Yes	Yes
Number of observations	800	800	800	800	800	800

Note: This table presents OLS regression estimates of household-level insurance uptake on economic preferences, based on double-bounded dichotomous choice questions. All specifications include distance to watercourse, covariates, urban community and flood exposure fixed effects. Standard errors are bootstrapped (1,000 replications). Asterisk (*), double asterisk (**), and triple asterisk (***) denote variables significant at 10 percent, 5 percent, and 1 percent, respectively.
Source: Household Survey on the Impact of Floods in Lima (2023).

Trust effects on insurance uptake

Trust has shown to be an important issue in shaping insurance take-up among farmers (Dercon, Gunning, and Zeitlin, 2011; Cole et al., 2013; Casaburi and Willis, 2018). For instance, if households believe the insurer is reliable and likely to pay claims, they are more likely to invest in insurance. This study posits that the level of trust citizens place in their government's intent and ability to manage flood risk significantly influences household flood preparedness.

Table 4.12 Trust impacts on extreme risk-averse households and insurance uptake

Model: OLS	Dependent variable: Insurance take-up							
	(1)		(2)		(3)		(4)	
Number of patient choices (in logs)	0.024	**	0.025	**	0.023	*	0.022	*
	(0.012)		(0.012)		(0.013)		(0.013)	
Extreme risk aversion (1,0)	-0.080	***	-0.081	***	-0.103	***	-0.176	***
	(0.028)		(0.029)		(0.036)		(0.034)	
Premium (in logs)	-0.098	***	-0.097	***	-0.098	***	-0.098	***
	(0.010)		(0.010)		(0.010)		(0.010)	
Trust in local government (1,0)	-0.015				-0.033			
	(0.029)				(0.034)			
Trust in local government flood plan (1,0)			-0.028				-0.104	***
			(0.029)				(0.033)	
Extreme risk aversion (1,0) x Trust in local government (1,0)					0.073			
					(0.062)			
Extreme risk aversion (1,0) x Trust in local government flood plan (1,0)							0.285	***
							(0.055)	
Distance to watercourse (in logs)	Yes		Yes		Yes		Yes	
Urban community FE	Yes		Yes		Yes		Yes	
Flood exposure FE	Yes		Yes		Yes		Yes	
Covariates	Yes		Yes		Yes		Yes	
Number of observations	800		800		800		800	

Note: This table presents OLS regression estimations of household-level insurance uptake on economic preferences, premium rates and trust in local government and flood plan. All sample corresponds to double-bounded dichotomous choice questions. All specifications include distance to watercourse, covariates, urban community and flood exposure fixed effects. Standard errors are bootstrapped (1,000 replications). Asterisk (*), double asterisk (**), and triple asterisk (***) denote variables significant at 10 percent, 5 percent, and 1 percent, respectively.
Source: Household Survey on the Impact of Floods in Lima (2023).

In columns (1) to (4) of Table 4.12, the results for economic preferences remain relatively consistent in magnitude and sign compared to the main estimates. Regressing trust in the local government (column 1) and trust in the local government's flood plan (column 2) reveals no significant effects on index-based insurance uptake. However, in column (4), I find a significant increase in demand for index-based insurance arising from the interaction between extreme risk aversion and trust in the local government's flood plan. A 1 SD increase in trust in the local government flood plan is associated with an 8.8 percentage points increase in demand for extremely risk-averse households.

Specifically, households that are highly risk-averse are more likely to purchase index-based insurance when they trust that the local government has a credible flood plan in place. This suggests that improving trust in the local government's capacity to manage floods in districts of Lurigancho–Chosica could influence preferences for financial risk management tools and increase insurance uptake among the most risk-averse households.

4.5.4 Heterogeneity analysis

This study also examines the effects of economic preferences on insurance demand among households based on their socioeconomic status and whether they are located within a flood zone.

In column (2) of Table 4.13, I observe that more patient and wealthier households are more likely to be insured against flood effects. High-income households experience less economic insecurity, so they do not need to focus as much on immediate consumption. Their wealth provides a buffer against short-term risks, allowing them to think more strategically about future rewards, such as insurance coverage.

In column (3), extremely risk-averse households are less likely to purchase insurance when located in high-risk flood areas. This could be explained by the fact that poorer households are usually found in areas exposed to floods and low amenities. Finally, insurance prices are consistently negatively associated with insurance uptake as shown in columns (1) through (4). These results emphasize that higher premiums deter households from purchasing insurance, regardless of their income level or proximity to flood-prone areas.

Regarding the time preference coefficients, in columns (3), (5) and (7) of Table 4.14, I find that low-income households with higher patience located within the flood zone are actually more likely to take up insurance provided by private companies. This could indicate a perception that private insurance offers better coverage or benefits for those households in flood-prone areas. In contrast, for the extreme risk aversion estimates, in columns (5), (7), (9), (10) and (11), I observe that low- and high-income households located in flood zone areas are more likely to take up insurance when offered by international organizations or NGOs, potentially due to subsidised premiums (perceived affordability or accessibility) or perceived trustworthiness and community-focused initiatives.

Table 4.13 Heterogeneity of the effects of economic preferences

Model: OLS	Dependent variable: Insurance uptake (1) Low income		(2) High income		(3) Within flood zone		(4) Outside flood zone	
Number of patient choices (in logs)	-0.017		0.078	**	0.023		0.029	
	(0.028)		(0.032)		(0.016)		(0.021)	
Extreme risk aversion (1,0)	-0.044		0.042		-0.085	**	-0.017	
	(0.064)		(0.087)		(0.035)		(0.051)	
Premium (in logs)	-0.085	***	-0.105	***	-0.098	***	-0.101	***
	(0.023)		(0.024)		(0.012)		(0.019)	
Distance to watercourse (in logs)	Yes		Yes		Yes		Yes	
Urban community FE	Yes		Yes		Yes		Yes	
Flood exposure FE	Yes		Yes		No		No	
Covariates	Yes		Yes		Yes		Yes	
Number of observations	204		196		558		242	

Note: This table presents OLS regression estimates of household-level insurance uptake on economic preferences, based on double-bounded dichotomous choice questions. All specifications include distance to watercourse, covariates, urban community and flood exposure fixed effects. Standard errors are bootstrapped (1,000 replications). Asterisk (*), double asterisk (**), and triple asterisk (***) denote variables significant at 10 percent, 5 percent, and 1 percent, respectively.
Source: Household Survey on the Impact of Floods in Lima (2023).

In general, it appears that households trust international organizations and NGOs more that private companies to provide insurance, especially among extremely risk-averse and low-income groups. These organizations could focus on further subsidising premiums or tailoring coverage for high-risk and low-income households to increase uptake, particularly when considering that 8.7% of the sample are extreme risk-averse households.

Table 4.14 Heterogeneity in the effects of economic preferences on trust in different agencies to provide insurance

	Dependent variable: Insurance uptake											
	Local government				Insurance company				International Organisation/NGO			
Model: OLS	(1) Low income	(2) High income	(3) Within flood zone	(4) Outside flood zone	(5) Low income	(6) High income	(7) Within flood zone	(8) Outside flood zone	(9) Low income	(10) High income	(11) Within flood area	(12) Outside flood zone
Number of household patient choices (in logs)	-0.045	0.034	-0.052***	0.013	0.059**	-0.048	0.052***	0.010	-0.026	0.014	-0.024	-0.031
	(0.035)	(0.031)	(0.018)	(0.026)	(0.030)	(0.042)	(0.018)	(0.028)	(0.031)	(0.045)	(0.019)	(0.025)
Extreme risk aversion (1,0)	-0.088	-0.049	-0.055	0.023	-0.124**	-0.119	-0.124***	0.054	0.155**	0.197*	0.157***	-0.066
	(0.078)	(0.055)	(0.039)	(0.057)	(0.059)	(0.096)	(0.035)	(0.068)	(0.066)	(0.109)	(0.045)	(0.069)
Premium (in logs)	0.007	0.018	0.024*	-0.008	-0.034*	-0.029	0.002	-0.010	0.017	-0.008	-0.027**	0.024
	(0.024)	(0.017)	(0.012)	(0.015)	(0.019)	(0.026)	(0.011)	(0.022)	(0.021)	(0.027)	(0.014)	(0.022)
Distance to watercourse (in logs)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Urban community FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Flood exposure FE	Yes	Yes	No	No	Yes	Yes	No	No	Yes	Yes	No	No
Covariates	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Number of observations	204	196	558	242	204	196	558	242	204	196	558	242

Note: This table presents OLS regression estimates of household-level insurance uptake on economic preferences, based on double-bounded dichotomous choice questions. All specifications include distance to watercourse, covariates, urban community and flood exposure fixed effects. Standard errors are bootstrapped (1,000 replications). Asterisk (*), double asterisk (**), and triple asterisk (***) denote variables significant at 10 percent, 5 percent, and 1 percent, respectively.
Source: Household Survey on the Impact of Floods in Lima (2023).

4.6 Policy implications

Participation in insurance schemes remains low in developing countries, so efforts should focus on boosting participation by providing information on the benefits of payouts, offering subsidies, and increasing trust. Insurance products function by transferring income across time: households make payments (such as premiums) today in exchange for compensation for potential future losses, such as those caused by floods. However, this transfer of income across time can reduce demand for insurance, especially among households with high time preferences (i.e., a preference for sooner, smaller rewards) and liquidity constraints (Baillon et al., 2022).

Building trust

The lack of trust in local governments among highly risk-averse households can be a significant barrier to increasing insurance uptake. Governments in developing countries should focus on building trust in their ability to manage and respond to local floods, particularly by involving urban communities in disaster risk management and flood response plans. Community engagement in disaster risk reduction plans could strengthen trust in the local government of Lurigancho-Chosica and encourage households to adopt higher levels of flood preparedness, such as purchasing index-based insurance. Empowering communities through participation in decision-making processes would increase the likelihood that households in flood-prone areas of Lima engage in flood risk mitigation measures and enhance overall community resilience for future disaster events.

Subsidising premiums

Subsidising insurance premiums could lower upfront costs, making insurance more accessible to flood-prone households, particularly those that are highly risk-averse and located in high-risk areas[11]. In general, insurance products priced below PEN 20 (USD 5.42) are more likely to be purchased; however, when prices exceed this threshold, the likelihood of purchase declines significantly among households (see Table 4.A.5). Moreover, the mean willingness to pay (MWTP) per household is 4.91% (PEN 11.15 or USD 3.02) of the monthly premium cost (PEN 227.27 or USD 61.61). Scaling this to the number of households

[11] Comparing with other disaster insurance schemes in Peru, a successful example of free disaster insurance is the Catastrophic Agricultural Insurance (SAC), the country's largest multiple-peril crop insurance program. For the 2023-2024 agricultural season, it had a budget of PEN 98 million (USD 26.6 million). SAC is funded by the Peruvian government through the Agricultural Guarantee and Insurance Fund (FOGASA). Its purpose is to protect the crops of nearly 320,000 small-scale farmers and reduce their vulnerability to disaster events.

exposed to floods (5,210 households), the total cost of providing index-based insurance in Lurigancho–Chosica is approximately PEN 14.21 million (USD 3.85 million), requiring a subsidy of nearly 95% (PEN 13.51 million or USD 3.66 million)[12]. This highlights the significant financial barrier that these households face in obtaining insurance (see Table 4.A.6).

4.7 Conclusion

Megacities in developing countries are often severely impacted by floods due to inadequate financing for resilient infrastructure and limited improvements in risk management. Index-based insurance has significant potential for enhancing flood resilience in these economies, particularly by helping vulnerable urban populations manage weather-related risks. This study presents causal evidence on three factors that affect the take-up of flood-based index-based insurance in Peru: level of patience (a proxy for time preference), risk aversion and insurance premium.

I collected data from experimental "games" (such as discount rate, risk-taking and bidding prices exercises) in three urban communities in Lurigancho-Chosica, a district within the megacity of Lima. This data includes bids on hypothetical index-based flood insurance and exogenous variations in time and risk preferences, as well as insurance prices. I control for household and geographic characteristics, flood impacts, mitigation measures, and the distance to the nearest weather station (as a proxy for basis risk).

The study finds that the uptake of index-based flood insurance increases with higher levels of patience. Impatient households are not only less likely to purchase insurance, but their strong preference for immediate benefits remains unchanged even as their spending power increases. Furthermore, liquidity constraints—experienced by 70% of households in the sample—may further deter these households from purchasing insurance. The waiting period between claim approval and fund disbursement also significantly influences the purchase and success of an index-based insurance scheme, as impatient households who lack access to affordable borrowing options may choose to forgo insurance altogether.

Conversely, demand for index-based insurance decreases at higher levels of risk aversion. Extremely risk-averse households, even those within flood-prone areas, are unwilling to pay for insurance—even at a significantly reduced premium. This behaviour is likely driven by a

[12] See Total subsidy estimation to implement insurance scheme in the Appendix.

lack of trust in local government and the insurance system to provide timely compensation. Instead, these households express a preference for international organizations or NGOs to manage insurance funds and deliver insurance products.

Regarding trust as a critical factor in shaping insurance uptake, results indicate that trust in the local government's ability to manage flood risk and provide flood insurance increases the likelihood of a household purchasing index-based insurance. Building this trust is essential for governments in developing countries, as it encourages households—particularly those that are highly risk-averse—to adopt stronger flood preparedness measures.

Enhanced trust should be paired with targeted subsidy policies to boost adoption among urban households in Lurigancho-Chosica. Premium reductions should prioritise households most in need, especially those located in flood-prone areas.

References

ACT Alliance 2012, <https://reliefweb.int/report/peru/peru-flash-mud-floods-lima-city-peru-per121>.

Almås, I, Cappelen, A, Salvanes, KG, Sorensen, EO, Tungodden, B 2016, 'Willingness to compete: family matters', *Management Science*, no. 62, pp. 2149–62.

Baillon, A, O'Donnell, O, Quimbo, K, Wilgenburg, KV 2022, Do time preferences explain low health insurance take-up?', Journal of Risk and Insurance, vol. 89, no. 4, pp. 951–983.

Barnes, ML, Wang, PC, Graham, NAJ, Guerrero, AM, Jasny, L, Lau, JJ, Sutcliffe, SR, Zamborain-Mason, J 2020, 'Social determinants of adaptive and transformative responses to climate change', *Nature Clinical Practice Endocrinology and Metabolism*, vol. 10, pp. 823–828.

Bauer, M, Chytilova, J, Morduch J 2012, 'Behavioural foundations of microcredit: experimental and survey evidence from rural India', *American Economic Review*, no. 102, pp. 1118–39.

Belissa, T, Bulte, E, Cecchi, F, Gangopadhyay, S, Lensink, R 2019, 'Liquidity constraints, informal institutions, and the adoption of weather insurance: a randomized controlled trial in Ethiopia', *Journal of Development Economics*, vol. 140, pp. 269–278.

Binswanger, H 1980, 'Attitudes toward risk: experimental measurement in rural India', *American Journal of Agricultural Economics*, vol. 62, no. 3, pp. 395–407.

Binswanger-Mkhize, HP 2012. Is there too much hype about index-based agricultural insurance? Journal of Development Studies, vol. 48, no. 2, pp. 187-200.

Brunner, MI, Swain, DL, Wood, RR, Willkofer, F, Done, JM, Gilleland, E, and Ludwig, R 2021, 'An Extremeness Threshold Determines the Regional Response of Floods to Changes in Rainfall Extremes', *Communications Earth and Environment*, vol. 2, no. 1.

Cao, W, Yang, Y, Huang, J, Sun, D, Liu, G 2020, 'Influential factors affecting protective coping behaviours of flood disaster: a case study in Shenzhen, China', *International Journal of Environmental Research and Public Health*, vol. 17, no. 16, 5945.

Casaburi, L, Willis, J 2018, ‘Time vs. state in insurance: experimental evidence from contract farming in Kenya’, *American Economic Review*, vol. 108, no. 12, pp. 3778–3813.

Chowdhury, S, Sutter, Zimmermann, KF 2022, ‘Economic preferences across generations and family clusters: a large-scale experiment in a developing country’, Journal of Political Economy, vol. 130, no 9, pp. 2361–2419.

Cole, S, Gin, X, Tobacman, J, Topalova, P, Townsend, R, Vickery, J 2013, ‘Barriers to household risk management: evidence from India’, American Economic Association, vol. 5, no. 1, pp. 104–135.

CRED Publication 2023, *Disasters in Numbers - A Significant Year of Disaster Impact*, <https://files.emdat.be/reports/2023_EMDAT_report.pdf>.

Dercon, S, Gunning, JW, Zeitlin, A 2011, ‘The demand for insurance under limited trust: evidence from a field experiment in Kenya, working paper, *Centre for the Study of African Economies*, University of Oxford.

Eckel, CC, Grossman, PJ 2002. ‘Sex differences and statistical stereotyping in attitudes toward financial risk’, *Evolution and Human Behaviour*, vol. 23 no. 4, pp. 281–295.

Ejeta, LT, Ardalan, A, Paton, D, Yaseri, M 2018, ‘Emotional and cognitive factors influencing flood preparedness in Dire Dawa town, Ethiopia’, *Natural Hazards*, vol. 93, no. 2, pp. 715–737.

Fan, J, Gijbels, I 1996, ‘Local polynomial modelling and its applications’, Open Journal of Statistics, vol.6, no. 6, pp. 1085-1097.

Giné, X., Townsend, R., Vickery, J., 2008. Patterns of rainfall insurance participation in rural India. World Bank Econ. Rev. 22 (3), 539–566.

Hanemann, M 1991, ‘Willingness to pay and willingness to accept: how much can they differ?’, *American Economic Review*, vol. 81, issue 3, pp. 635–47.

Hill, RV, Robles, M, Ceballos, F 2016, ‘Demand for a simple weather insurance product in India: theory and evidence’, *American Journal of Agricultural Economics*, vol. 98, pp. 1250–1270.

INEI - National Institute of Statistics and Information of Peru, INEI Bulletin 2023 <https://m.inei.gob.pe/media/MenuRecursivo/noticias/nota-de-prensa-no-104-2023-inei.pdf>

Janzen, SA, Carter, MR 2013, 'After the drought: the impact of microinsurance on consumption smoothing and asset protection', *National Bureau of Economic Research*.

Kazianga, H, Udry, C 2006, 'Consumption smoothing? livestock, insurance and drought in rural Burkina Faso', *Journal Development Economics*, vol. 79, no. 2, pp. 413–446.

Kundu, A 2022, 'Do risk and time preference explain household's demand for microinsurance? a lab in the field approach', *SSRN*, <https://ssrn.com/abstract=4113198>.

Liu, Y, Chen, K, Hill, RV 2020, 'Delayed premium payment, insurance adoption, and household investment in rural China', *American Journal of Agricultural Economics*, vol. 102, no.4, pp. 1177–1197.

Malik, RPS, Amarnath, G 2021, 'Economics of index-based flood insurance (IBFI): scenario analysis and stakeholder perspectives from South Asia', *International Water Management Institute*.

Nadaraya, EA 1965, 'On non-parametric estimates of density functions and regression curves', *Theory of Probability and its Applications*, vol. 10, pp. 186–190.

Rentschler, J, Salhab, M and Jafino BA, "Flood Exposure and Poverty in 188 Countries," *Nature Communications*, 2022, 13.

Shin, S, Magnan, N, Mullally, C Janzen, S 2022, 'Demand for weather index insurance among smallholder farmers under prospect theory', *Journal of Economic Behaviour and Organization*, vol. 202, pp. 82-104.

Soane, E, Schubert, I, Challenor, P, Lunn, R, Narendran, S, Pollard, S 2010, 'Flood perception and mitigation: the role of severity, agency, and experience in the purchase of flood protection, and the communication of flood information', *Environment and Planning a: Economy and Space*, vol. 42, no. 12, pp. 3023–3038.

Stigler, GJ, Becker, GS 1977, 'De gustibus non est disputandum', *American Economic Review*, vol. 67, pp. 76–90.

Ullah, F, Shah, SAA, Saqib, SE, Yaseen, M, Haider, MS 2021, ‘Households’ flood vulnerability and adaptation: empirical evidence from mountainous regions of Pakistan’, *International Journal of Disaster Risk Reduction*, vol. 52, 101967.

Watson, GS 1964, ‘Smooth regression analysis’, *Sankhya Series A*, vol. 26, pp. 359–372.

Appendix

Section A. Sample size and proportion to the sample assigned to each group

The study uses a simple random sampling formula for proportions in finite populations to calculate the sample size. The formula used is:

$$n = \frac{N * Z^2 * p * q}{(N-1) * e^2 + Z^2 * p * q}$$

where n is the sample size, N is the total number of households, Z is the Z-score of the desired confidence level (1.96 for 95%), p and q are the probabilities of the event occurring and not occurring (set to 0.5 each to maximise variability), and e is the margin of error (set at 5%). This formula incorporates a finite population adjustment factor, reducing the sample size compared to infinite population scenarios. For this study, with a total population of 3,239 households, the optimal sample size was calculated to be 363 households. During the survey, a total of 404 household were included in the sample. This ensures sufficient statistical power and precision to estimate population parameters within the specified confidence and error margins.

Section B. Additional tables and figures

Table 4.A.1 Distribution of sampled households by choice set sequence

Choice set	Sequence	Number of sample	Percentage of sample
1	(1, 2, 3)	71	17.57%
2	(1, 3, 2)	68	16.83%
3	(2, 3, 1)	57	14.11%
4	(2, 1, 3)	61	15.10%
5	(3, 1, 2)	77	19.06%
6	(3, 2, 1)	70	17.33%
Total		**404**	**100%**

Source: Household Survey on the Impact of Floods in Lima (2023).

Table 4.A.2 Distribution of sampled households by the degree of risk aversion

Risk aversion levels	Real payoffs	Number of sample	Percentage of sample
Neutral-to-negative	PEN 0.00 - PEN 5.00	47	11.63%
Slight-to-neutral	PEN 0.40 - PEN 4.00	38	9.41%
Moderate	PEN 0.80 - PEN 3.50	56	13.86%
Intermediate	PEN 1.20 - PEN 3.00	82	20.30%
Severe	PEN 1.60 - PEN 2.50	62	15.35%
Extreme	PEN 2.00 - PEN 2.00	119	29.46%
Total		**404**	**100%**

Source: Household Survey on the Impact of Floods in Lima (2023).

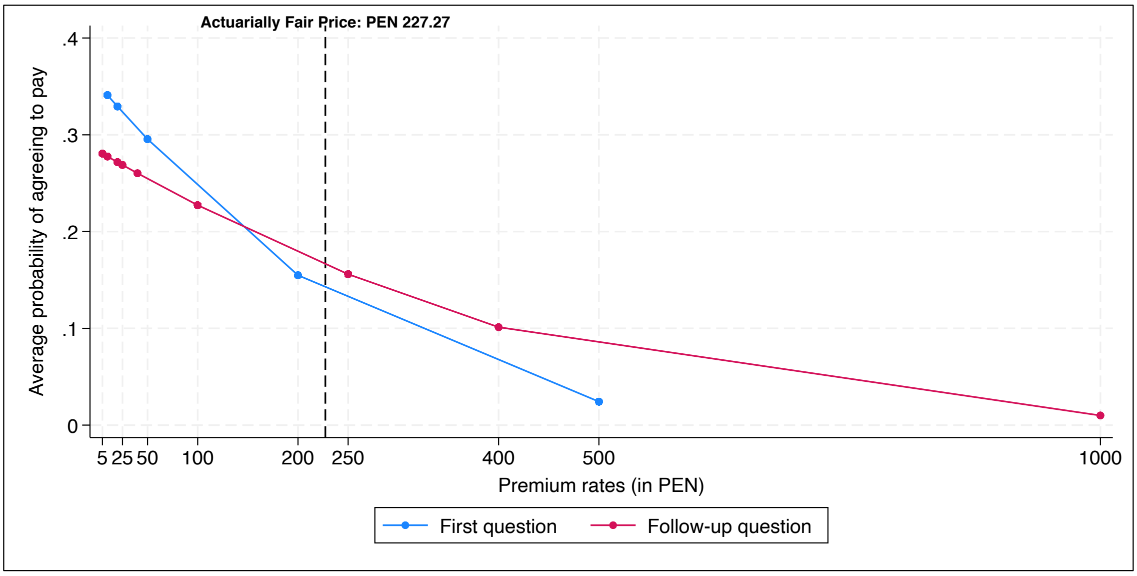


Figure 4.A.1 Marginal effect of household's WTP for index-based flood insurance at different bid prices

Note: This figure shows the probability of insurance purchase for the first question and the follow-up question under a dichotomous choice method.
Source: Household Survey on the Impact of Floods in Lima (2023).

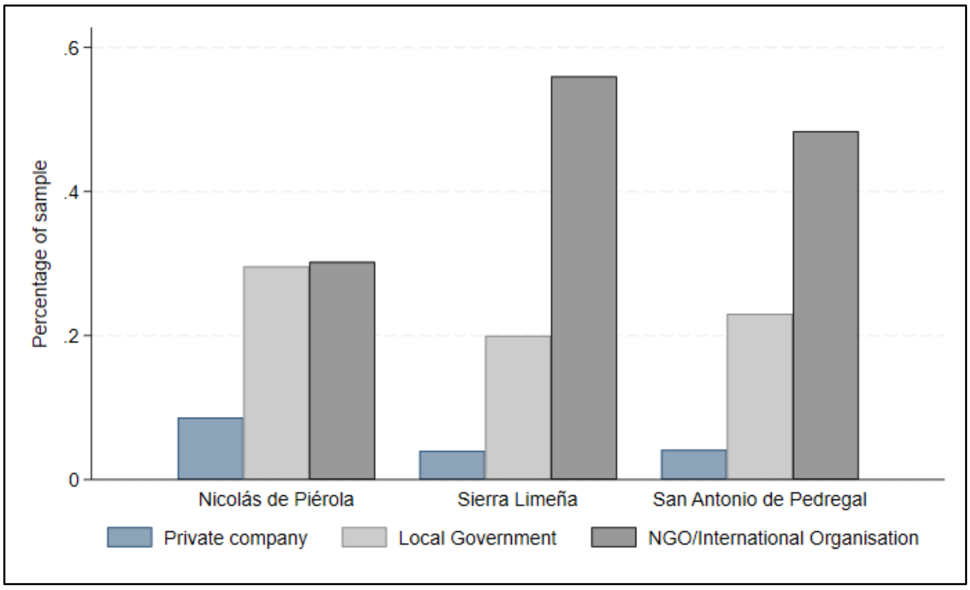


Figure 4.A.2 Level of trust in different types of insurers across urban communities

Note: This figure shows the percentage of respondents with the highest level of trust in one of the following insurers (private company, local government, or NGO/International Organisation) across the urban communities of Nicolás de Piérola, Sierra Limeña, and San Antonio de Pedregal. For each insurer type, a dummy variable is assigned a value of one if the respondent selects that insurer, and zero otherwise.
Source: Household Survey on the Impact of Floods in Lima (2023).

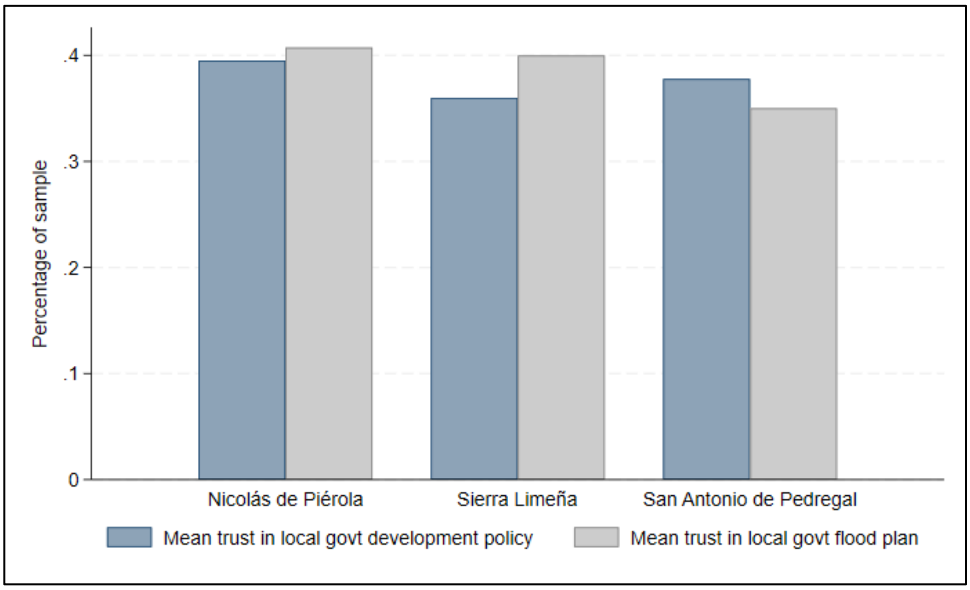


Figure 4.A.3 Level of trust in local government development policy and in a flood mitigation plan across the three urban communities

Note: This figure shows the percentage of respondents who expressed trust in the local government's development policy (infrastructure, education, and health) and its flood mitigation plan across the three urban communities of Nicolás de Piérola, Sierra Limeña, and San Antonio de Pedregal.
Source: Household Survey on the Impact of Floods in Lima (2023).

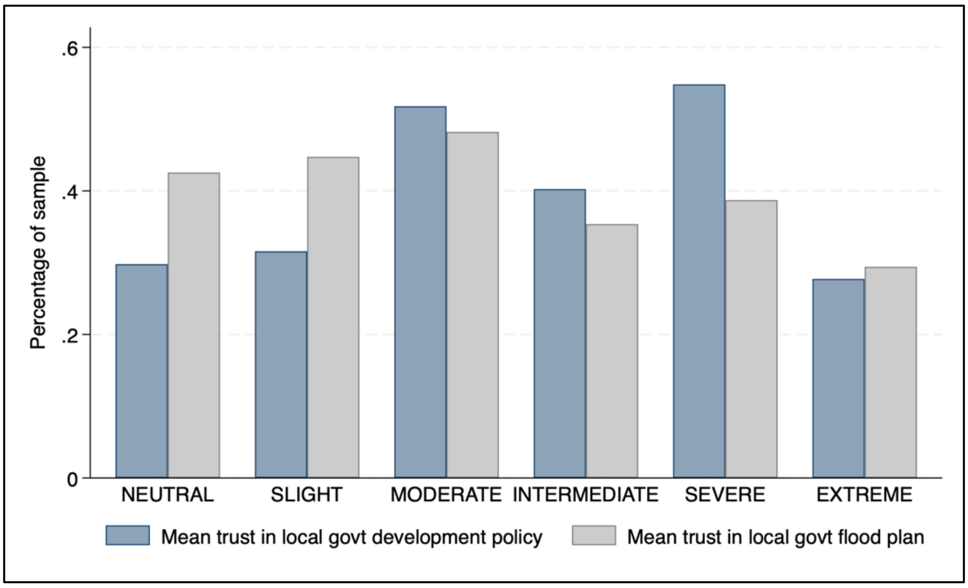


Figure 4.A.4 Level of trust in local government development policy and in a flood mitigation plan across different measures of risk aversion

Note: This figure shows the percentage of respondents with the highest level of trust in one of the following insurers (private company, local government, or NGO/International Organisation) across the urban communities of Nicolás de Piérola, Sierra Limeña, and San Antonio de Pedregal. For each insurer type, a dummy variable is assigned a value of one if the respondent selects that insurer, and zero otherwise.
Source: Household Survey on the Impact of Floods in Lima (2023).

Table 4.A.3 Exogeneity tests: OLS regressions before PSM

Sample: **Dependent variable:**	**Number of patient choices (in logs)**			**Extreme aversion (1,0)**			**Premium rates (in logs)**	
Electricity (1,0)	-0.0086	(0.0065)		0.0076	(0.0146)		-0.0050	(0.0044)
Water (1,0)	-0.0061	(0.0180)		-0.0134	(0.0405)		-0.0091	(0.0121)
Sewage (1,0)	-0.0183	(0.0123)		0.0369	(0.0277)		0.0115	(0.0083)
Toilet (1,0)	-0.0029	(0.0043)		0.0128	(0.0096)		-0.0006	(0.0029)
Floor (1,0)	-0.0002	(0.0073)		-0.0225	(0.0164)		-0.0062	(0.0049)
Storey (number)	0.0056	(0.0385)		-0.1442	(0.0864)	*	-0.0311	(0.0259)
Have land title (1,0)	0.0018	(0.0227)		-0.0197	(0.0510)		-0.0127	(0.0153)
Female (1,0)	-0.0064	(0.0233)		0.0921	(0.0522)	*	0.0247	(0.0156)
Age (in logs)	-0.0124	(0.0157)		0.0018	(0.0354)		-0.0159	(0.0106)
Married (1,0)	0.0044	(0.0228)		0.0074	(0.0514)		-0.0226	(0.0153)
Education (years)	-0.1732	(0.2079)		-0.1687	(0.4684)		-0.0696	(0.1401)
Household size (in logs)	0.0187	(0.0186)		0.0072	(0.0418)		-0.0022	(0.0125)
Self-/employed last week (1,0)	0.0118	(0.0209)		-0.0004	(0.0470)		0.0078	(0.0140)
Household expenditure (in logs)	-0.0511	(0.0613)		-0.1849	(0.1379)		0.0458	(0.0413)
Public health insurance (1,0)	-0.0043	(0.0200)		-0.0064	(0.0449)		0.0205	(0.0134)
Flood experience 2015-2023 (1,0)	-0.0057	(0.0185)		0.0287	(0.0416)		0.0146	(0.0124)
Flood expenditure 2017 (in logs)	0.0577	(0.0756)		-0.0777	(0.1193)		0.0181	(0.0357)
Floods in last 10 years (number)	0.0045	(0.0530)		0.0393	(0.1702)		0.0496	(0.0509)
Household mitigation (1,0)	0.0107	(0.0239)		-0.0411	(0.0538)		0.0031	(0.0161)
Distance to weather station (km)	-0.0144	(0.0087)	*	0.0174	(0.0196)		-0.0055	(0.0058)
Distance to barrier (in logs)	-0.0281	(0.0197)		0.0028	(0.0445)		-0.0069	(0.0133)
Elevation (in logs)	0.0001	(0.0061)		0.0121	(0.0137)		0.0031	(0.0041)
Distance to watercourse (in logs)	Yes			Yes			Yes	
Urban community FE	Yes			Yes			Yes	
Number of observations (N)	404			404			404	
Test of joint significance	F-stat: 0.72 (p-value: 0.8244)			F-stat: 0.80 (p-value: 0.7259)			F-stat: 0.87 (p-value: 0.6355)	

Source: Household Survey on the Impact of Floods in Lima (2023).

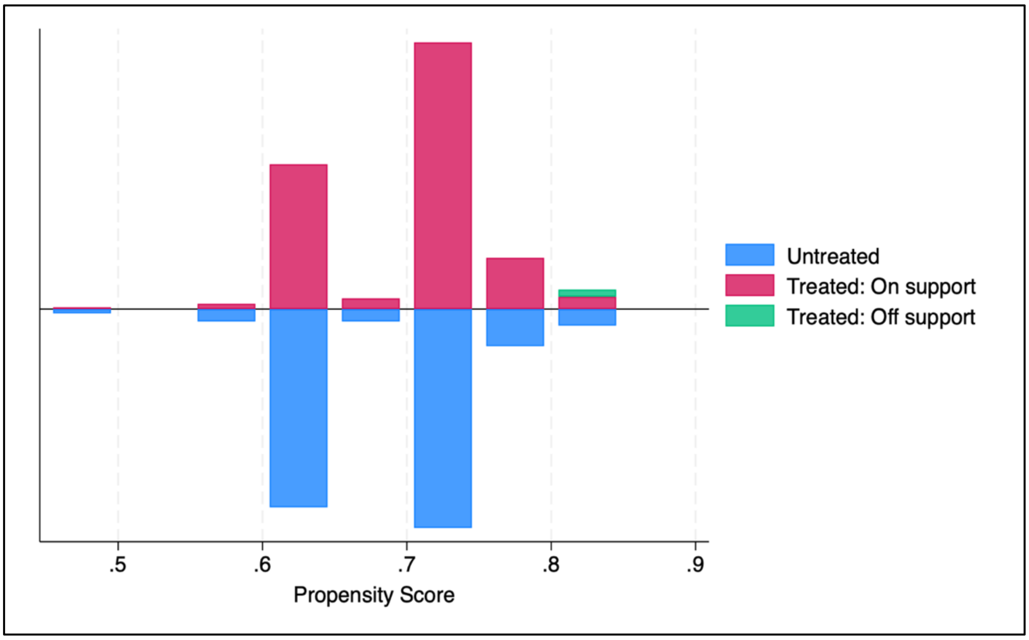


Figure 4.A.5 Propensity score matching graph

Source: Household Survey on the Impact of Floods in Lima (2023).

Table 4.A.4 Balance tests between flood-free and flood-prone urban communities before PSM

	(1)	(2)	(3)	(4)	(5)	
	Flood-free group	Flood-prone group	Single difference	N	p-value	
A. Variables of interest						
Number of patient choices (in logs)	1.294	1.230	0.064	404	0.720	
Extreme aversion (1,0)	0.256	0.311	-0.055	404	0.270	
Premium - First question (in logs)	5.000	4.508	0.492	404	0.003	***
B. Covariates						
Electricity (1,0)	0.975	0.986	-0.011	404	0.453	
Water (1,0)	0.174	0.159	0.015	404	0.718	
Sewage (1,0)	0.950	0.926	0.025	404	0.365	
Toilet (1,0)	0.992	0.993	-0.001	404	0.898	
Floor (1,0)	0.967	0.982	-0.015	404	0.338	
Storey (number)	1.802	1.686	0.116	404	0.188	
Have land title (1,0)	0.397	0.318	0.079	404	0.127	
Female (1,0)	0.694	0.671	0.023	404	0.654	
Age (in logs)	3.871	3.906	-0.034	404	0.321	
Married (1,0)	0.314	0.293	0.021	404	0.677	
Education (years)	14.008	13.940	0.068	404	0.882	
Household size (in logs)	1.582	1.601	-0.019	404	0.638	
Self-/employed last week (1,0)	0.273	0.205	0.068	404	0.136	
Household expenditure (in logs)	6.959	6.774	0.185	404	0.168	
Public health insurance (1,0)	0.826	0.788	0.038	404	0.378	
Distance to weather station (km)	2.087	2.084	0.003	404	0.964	
Distance to barrier (in logs)	7.018	7.071	-0.053	404	0.244	
Elevation (in logs)	6.837	6.847	-0.010	404	0.474	

Note: This table presents the summary statistics (standard errors in brackets) for the sample used in this study, along with the p-values for differences in means between the randomly assigned flood-free or flood-prone groups.
Source: Household Survey on the Impact of Floods in Lima (2023).

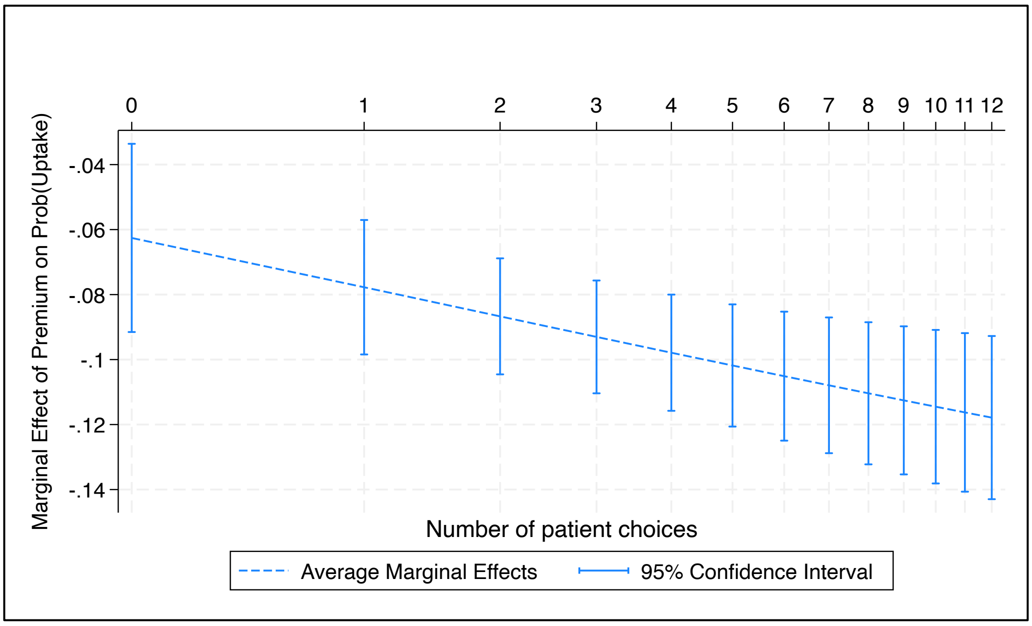


Figure 4.A.6 Price sensitivity of insurance demand and different level of patience

Notes: This figure shows the probability of insurance purchase against different levels of patience. The dashed line plots average marginal effects of price on insurance uptake. The solid lines plot 95 percent confidence bands around the average marginal effect point estimates. The figure stems from the estimation of an OLS regression model of household-level insurance uptake, levels of patience, and the interaction between the two. This specification includes distance to watercourse, covariates, as well as flood exposure and urban community fixed effects.
Source: Household Survey on the Impact of Floods in Lima (2023).

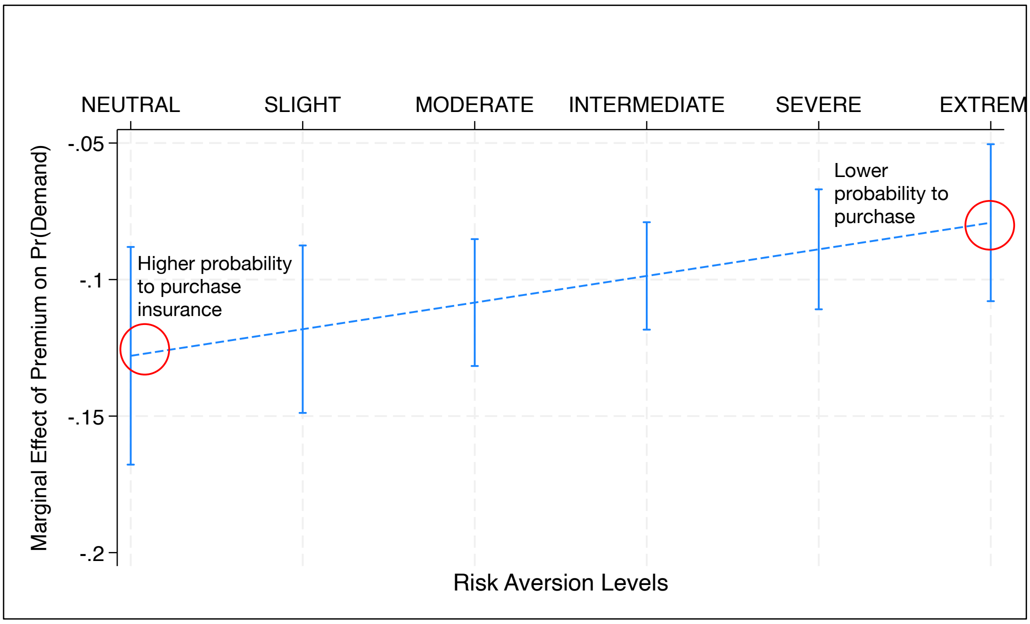


Figure 4.A.7 Probability of insurance purchase across levels of risk aversion

Notes: This figure shows the probability of insurance purchase against different levels of risk aversion. The dashed line plots average marginal effects of price on insurance uptake. The solid lines plot 95 percent confidence bands around the average marginal effect point estimates. The figure stems from the estimation of an OLS regression model of household-level insurance uptake, levels of risk aversion, and the interaction between the two. This specification includes distance to watercourse, covariates, as well as flood exposure and urban community fixed effects.
Source: Household Survey on the Impact of Floods in Lima (2023).

Table 4.A.5 Premium rates and demand for index-based flood insurance

	Dependent variable: Insurance demand					
Model: OLS	(1) All sample		(2) 1st question		(3) 2nd question	
Premium rates						
PEN 10	0.149	*			0.050	
	(0.083)				(0.105)	
PEN 20	-0.010		-0.413	***	0.307	***
	(0.087)		(0.078)		(0.110)	
PEN 25	-0.118				-0.109	
	(0.089)				(0.096)	
PEN 40	0.264				0.264	
	(0.249)				(0.247)	
PEN 50	-0.132		-0.350	***		
	(0.086)		(0.077)			
PEN 100	-0.213	***			-0.212	**
	(0.078)				(0.087)	
PEN 200	-0.236	***	-0.445	***		
	(0.078)		(0.066)			
PEN 250	-0.279	***			-0.280	***
	(0.072)				(0.077)	
PEN 400	(0.078)				(0.037)	
	(0.251)				(0.259)	
PEN 500	(0.316)	***	(0.518)	***		
	(0.071)		(0.059)			
PEN 1,000	0.293				0.344	
	(0.278)				(0.315)	
Number of patient choices (in logs)	0.023	*	0.003		0.042	**
	(0.013)		(0.017)		(0.019)	
Extreme risk aversion (1,0)	-0.075	***	-0.068	*	-0.056	
	(0.026)		(0.036)		(0.041)	
Distance to watercourse (in logs)	Yes		Yes		Yes	
Valley FE	Yes		Yes		Yes	
Flood exposure FE	Yes		Yes		Yes	
Covariates	Yes		Yes		Yes	
Number of observations	800		400		400	

Note: This table presents OLS regression estimates of household-level insurance uptake on economic preferences, based on double-bounded dichotomous choice questions. Premium rates come from the five price levels and their follow-up (discounted or increased) price under a double-bounded dichotomous choice. All price levels are dummy variables, except for PEN 5 (USD 1.36) which is omitted in the table. Average marginal effects are reported. All sample corresponds to double-bounded dichotomous choice questions in column (1), and the first and follow-up questions in columns (2) and (3), respectively. Standard errors are bootstrapped (1,000 replications) and reported in parentheses. Asterisk (*), double asterisk (**), and triple asterisk (***) denote variables significant at 10 percent, 5 percent, and 1 percent, respectively.
Source: Household Survey on the Impact of Floods in Lima (2023).

Total subsidy estimation to implement insurance scheme

In Table A.6, I observe that people's monthly WTP for index-based insurance is PEN 11.15 (USD 3.02), which represents 4.91% of the total insurance premium. To obtain this mean WTP, I used the estimated premium (in logs) and the constant from the OLS model with the following formula:

$$\text{Mean monthly WTP} = -\frac{\ln(1 + exp\{premim(in\ logs)\})}{constant}$$

If the local government plans to implement an index-based insurance scheme in the district of Lurigancho-Chosica, an annual subsidy of PEN 13.51 million (USD 3.66 million) would be required. This subsidy amounts to 0.003% of Lima's Gross Domestic Product in 2023.

Table 4.A.6 Cost of implementing index-based flood insurance

Model: OLS	Dependent variable: Insurance demand	
	(1)	
Premium (in logs)	-0.0924	***
	(0.0101)	
Constant	0.5894	
	(0.1272)	
Distance to watercourse (in logs)	Yes	
Urban community FE	Yes	
Flood exposure FE	Yes	
Covariates	No	
Number of observations	800	

Monthly premium	PEN 227.27 (USD 61.61)
Mean WTP	PEN 11.15 (USD 3.02)
Number of households[1/]	60,883
Number of households exposed to floods[2/]	5,210
Annual cost (million)	PEN 14.21 (USD 3.85)
Annual subsidy (million)	PEN 13.51 (USD 3.66)
Lima GDP (million)	PEN 430,791 (USD 116,777.17)
Share of Lima GDP (%)	0.003%

1/ Census 2017
2/ The National Water Authority of Peru Report (2017)
Source: Household Survey on the Impact of Floods in Lima (2023).

Chapter 5
Conclusion

5.1 Thesis highlights

Flooding is the most prevalent natural hazard that has significantly affected people around the world (Rentschler et al., 2020). Record-breaking floods, caused by increased rainfall due to climate change, in addition to rapid urbanisation and poor land management, pose new challenges, particularly for those who live in developing countries. Indonesia and Peru have faced severe consequences from flooding, with Jakarta, as well as the northern region and Lima, respectively, experiencing widespread disruption to people's livelihoods. Nevertheless, index insurance offers an innovative risk management solution for communities vulnerable to weather-related disasters. This risk transfer product shows significant promise for promoting sustainable development not only in the rural areas, as documented in literature, but also in enhancing resilience to flood shocks in megacities of developing countries.

This thesis comprises three essays evaluating the economic impacts of floods in northern Peru, and examining the role of basis risk, economic preferences, and trust in the implementation of index insurance as a risk transfer mechanism in the megacities of Jakarta and Lima. Chapter 2 tests theoretical predictions derived from an Expected Utility Model tailored to index-based flood insurance. Using Jakarta, Indonesia, as a case study, the chapter examines how basis risk, premium, and risk aversion influence demand for the insurance product. To achieve this, a household survey was conducted in early 2018 across the capital of Indonesia. The survey interviewed 836 households and, using a hypothetical index insurance product with basis risk, collected information on risk attitudes, experience with floods, and insurance demand through a multiple price list with price increase or discount based on their first response. The study uses a probit approach as the main specification, leveraging the distance to the floodgate – used as a proxy for basis risk – as a plausibly exogenous variable, after controlling for distance of the surveyed house to the nearest river. It assumes that residential location choices, whether near or far from the nearest river, are primarily driven by income, housing affordability, and access to basic services rather than proximity to the floodgate station. Other key variables, such as the premium and extreme risk aversion, are considered plausibly exogenous. The study explores the characteristics of demand for index-based flood insurance to cover household losses from yearly flooding in Jakarta.

Chapter 3 examines the economic impacts of flood shocks on household welfare and poverty resulting from the coastal El Niño phenomenon that affected northern Peru in 2017. To identify the affected districts, the study constructs a damage index using multiple remote sensing datasets and reports from the Peruvian government. This index serves as a proxy for local economic impact, providing a measure of damage severity across the region. A diff-in-diff event study design is employed to estimate the causal effects of this climatic shock on income per capita, expenditure per capita, and poverty levels before and after the floods. To address potential endogeneity arising from selectivity or sorting issues, propensity score matching (PSM) is applied. The identification strategy assumes that flooding is quasi-random, after controlling for the distance between surveyed households and the nearest river. Under this framework, the parallel trends and no-anticipation assumptions hold. To estimate the impact, I use secondary data from the ENAHO survey, which includes information on income, expenditure, and related variables, as well as household and dwelling characteristics. The analysis covers a total of 308 households across 57 districts in the northern region.

Chapter 4 explores the role of economic preferences and trust in the adoption of a hypothetical index insurance product provided by the local government in flood-prone areas on the eastern side of Lima, Peru. Primary data from a lab-in-the-field experiment with 404 households in three urban communities include experimental games to elicit patience—as a proxy for time preference—and risk aversion, as well as a multiple price list, to determine whether households are likely to purchase the hypothetical insurance product. Trust measures, on the other hand, are obtained by assessing social capital and institutional trust. In addition, the experiment collects information on socio-demographic characteristics, flood experience, and willingness to pay for index insurance.

The identification strategy assumes that economic preferences are exogenous after controlling for the distance between the surveyed house and the nearest river. The study employs an OLS approach as the main specification. Time preference is measured by the number of times a respondent chooses option 2, which offers a larger but delayed payoff. A higher number of such choices indicates greater patience, making the respondent more likely to purchase insurance. Risk preference is assessed through six investment options, each reflecting different levels of risk aversion. A household is classified as risk-seeking if they choose the riskiest option (1: PEN 0.00 – PEN 5.00) and as extremely risk-averse if they select the safest option (6: PEN 2.00 – PEN 2.00). Extreme risk aversion is associated with a higher willingness to purchase insurance.

5.2 General findings

In general, this thesis finds that flooding negatively affects household welfare by reducing income, expenditure per capita, and financial assets such as savings. Households with lower socioeconomic status are more vulnerable to flood damage and recover more slowly, especially in urban areas. This further entrenches poverty in affected districts. A possible explanation is the decline in the number of household members at work, as they lose their jobs and/or face wage reductions following flood events. However, donated goods such as food and clothing, provided by the private and public sectors, help reduce the immediate financial burden, allowing households to smooth their total consumption.

One potential countermeasure to mitigate these adverse effects is the development of index-based insurance, a risk management tool that shows significant promise for promoting sustainable development and resilience to weather shocks such as floods. Index-based insurance works by tying payouts directly to an objective, measurable index that correlates with losses caused by flooding. Fieldwork findings indicate that households in the megacities of Jakarta and Lima exhibit low demand for index-based insurance, at 9.9% and 11.4%, respectively. Nevertheless, opportunities exist to implement an index insurance scheme by targeting subsidies to households with higher price sensitivity and risk aversion, building trust in local government, and investing in monitoring infrastructure.

5.1.1 Specific findings

This section outlines the specific findings of the three studies in greater detail. In the first study, results indicate that demand for hypothetical index-based flood insurance in Jakarta decreases as basis risk, premiums, and risk aversion increase. On average, households offered insurance referenced to a nearby floodgate station—within 5 km—are more likely to purchase it, believing the station's floodwater level closely approximates their home's. This is supported by variations in price elasticity of demand, which decreases as households are closer to the reference floodgate (lower basis risk) and increases as they are farther away (higher basis risk). Moreover, demand declines at higher levels of risk aversion due to imperfect coverage, even with low basis risk. Conversely, households with lower extreme risk aversion exhibit higher price sensitivity, leading to greater willingness to purchase insurance.

The second study, focused on the impact of the 2017 coastal El Niño floods in northern Peru, shows that households in affected districts experience a 32.5% and 29.9% reduction in income and expenditure per capita, respectively. They are also 41.74 percentage points more likely to fall into poverty than those in unaffected districts. These adverse flood effects are more evident in urban than rural areas. The dynamic treatment effects persist after the disaster (2017–2019) for expenditure per capita and poverty, while income per capita recovers after two years (2017 and 2018). The results remain similar in magnitude and sign whether using the flood damage index or constructing the index with a different year of night-time lights, demonstrating that the treatment variable is as good as random assigned. Furthermore, the negative impact on household well-being appears to stem from a reduction in employed family members and declining wages. In response, households rely on financial assets such as savings and public or private donations of food and clothing to sustain total consumption following the disaster.

The third study finds that economic preferences influence index flood insurance uptake in Lima, increasing by 2.4 percentage points with higher levels of patience but decreasing by 3.6 percentage points and 8.2% with higher levels of risk aversion and premium, respectively. Regarding time preference, the claim disbursement period is crucial to the success of the insurance scheme among impatient households. For risk preference, the least risk-averse households are willing to pay a maximum of PEN 50 per unit for index insurance, whereas extremely risk-averse households are willing to pay PEN 0 per unit. A potential reason for reluctance to purchase insurance is a lack of trust in local government flood mitigation efforts among highly risk-averse households. While patient households prefer private companies as insurance providers—recognising the benefits of future payouts if a flood occurs—extremely risk-averse households tend to place greater trust in NGOs, which they perceive as more trustworthy and community-focused. Finally, a one-standard-deviation increase in trust in the local government's flood plan is associated with an 8.8 percentage point increase in demand among extremely risk-averse households.

A notable finding in this thesis is that the estimated prevalence of extreme risk aversion appears similar across households in the two study sites—approximately 31% in Jakarta and 30% in Lima. This consistency, observed through Binswanger-style lottery choices, suggests that people's underlying attitudes toward risk appear remarkably similar even across communities living in different parts of the world, with different flood experiences and institutional environments, yet sharing specific characteristics such as being low-income and located in flood-prone areas in the urban context of developing countries. A similar pattern emerges in the relationship between risk aversion and demand for insurance: extremely risk-averse households are unwilling to buy index insurance in both settings, though the mechanisms differ—basis risk in Jakarta and trust in the government flood plan in Lima. Taken together, these studies contribute to the external validity of behavioural, particularly risk preferences, in shaping the adoption of index-based flood insurance in the urban context.

On the other hand, the structural context—including flooding frequency, basis risk, trust in institutions, and pricing—may still drive heterogeneity in index insurance uptake between the two sites. In Jakarta, low demand for index insurance is likely driven by concerns over basis risk, due to the perceived unreliability of the index—linked to the lack of public infrastructure (i.e., a sparse network of floodgates)—even though the probability of flooding is around 20%. In contrast, in Lima, there is little trust in the government's ability to manage floods among households. As a result, flood insurance uptake remains low, despite a flood probability of 27%. In terms of pricing, the MWTP differs in each city—USD 0.57 in Jakarta and USD 3.02 in Lima; however, both represent approximately 4.9% of the total insurance premium, with subsidies considered vital to cover the upfront costs for households living in high flood-risk areas and who are extremely risk averse.

Therefore, these findings underscore the importance of tailoring index insurance design to the local institutional and environmental context, with risk preferences among households being an important component in generating effective policy responses for flood preparedness.

5.2 Policy insights and recommendations

The findings of this thesis have provided an assessment of the economic impact of flood shocks on household welfare and poverty, and evidence on the factors that influence demand for index-based flood insurance to cover losses affecting urban communities in the developing countries. These findings can help governments, private actors, NGOs, and other stakeholders design policies that efficiently allocate financial resources for flood disaster response and recovery. In addition, they can support the prioritisation of mitigation and adaptation strategies to reduce risk and impact while enhancing financial protection against flood-related losses, which are expected to increase in frequency and intensity due to climate change.

Efforts should be made to encourage households to engage in private flood preparedness measures, such as flood insurance, in urban areas of developing countries that are prone to devastating flooding. While empirical studies recognise the importance of these measures—given the limited capacity of city governments to implement comprehensive flood management strategies—the expansion of their role needs to be brought to the forefront. Chapter 2 suggests that investing in more floodgates could be effective in encouraging future demand for the index-based insurance product as a flood preparedness tool. New stations would monitor and collect data on river water levels for index references, particularly in the West and South regions, which could help reduce basis risk when flood indices fail to accurately reflect predicted losses. More broadly, investments in infrastructure that support data collection (such as river flow and precipitation levels) are essential for producing accurate and credible flood indices for insurance design (with low basis risk), while also reducing asymmetric information in the insurance market. This policy could be complemented by trust-building strategies that improve communication about potential flood-related risks households may face, encourage discussions on coping mechanisms, and provide information on the benefits of adopting index insurance. Strengthening both data infrastructure and trust in the product may increase households' willingness in Jakarta to adopt index insurance as a mechanism for flood preparedness.

On the other hand, as demonstrated in Chapter 4, a lack of trust among urban households in their local government's ability to manage flood impacts reduces the adoption of mitigation measures, such as purchasing index-based flood insurance, in Lima, Peru. A more inclusive and participatory approach to disaster risk management – involving vulnerable urban communities, including those with higher risk aversion – could encourage demand for index-based insurance. Strengthening preparedness and response mechanisms through collaborative efforts, such as engaging communities in the government's decision-making process for response plans, can help build trust in flood management capabilities. This policy may also involve third-party institutions, such as NGOs or universities, to provide education and training on basic concepts of risk and insurance, supporting the use of risk transfer instruments in underdeveloped insurance markets. In turn, this approach may encourage communities to respond positively by purchasing the insurance product as a preparedness measure.

Findings from research in Jakarta and Lima provide strong statistical evidence of a downward-sloping demand curve: as price increases, demand decreases, with marginal drops being particularly pronounced at lower price levels. This suggests that substantial subsidies may be necessary to stimulate demand, especially if targeted at poor and vulnerable households. However, when the benefits of investing in infrastructure for data collection are shared across multiple insurance contracts, each policyholder's contribution to the costs becomes relatively small. Such risk-sharing contracts enable insurance products to scale while lowering per-household costs over time. As coverage expands, investment costs are spread more broadly, reducing financial barriers to adoption.

Drawing on the impacts of the coastal El Niño floods on household welfare and poverty in the northern region of Peru, as discussed in Chapter 3, and considering the size of the estimated damage and the most affected areas, Peruvian disaster risk management agencies could allocate fiscal transfers from the central to local governments. This would facilitate the provision of emergency services and disaster relief, ensuring a more efficient response and the deployment of resources where they are most needed. In addition, the findings provide empirical evidence of average household income loss due to flood shocks, which can help determine the recommended amount of disaster insurance that could encourage the development of a risk transfer market, providing coverage in affected areas if a similar event occurs in the near future. Such disaster insurance could support vulnerable communities in adapting to and coping with extreme flood-related events by promoting risk-sharing mechanisms and self-insurance options.

5.3 Limitations

There are some potential limitations to the findings of this thesis. First, in Chapter 2, the variable distance to the floodgate, used as a proxy for basis risk, may still present endogeneity issues, potentially leading to biased findings. Regarding Chapter 4, several studies provide evidence that natural shocks can alter the risk and time preferences of affected households (see Page et al., 2012; Chantarat et al., 2019, for flood impacts on risk and time preferences in Australia and Cambodia). To address endogeneity concerns, the index-based flood insurance studies apply the Oster test to assess whether the results are driven by unobservable factors. Second, the pre-treatment analysis in Chapter 3 was limited by the five-year panel dataset used. In the diff-in-diff event study design, a pre-treatment period of only two years (2015–2016) restricts the application of pre-trend test procedures—such as those proposed by Roth (2022) and Rambachan and Roth (2023)—to support the validity of the parallel trends assumption.

5.4 Avenues for future research

There is scope for future research to identify key areas and strategies for mitigating the impact of flood shocks on household welfare. One potential avenue is exploring the optimal placement of floodgate stations to maximise insurance take-up and minimise basis risk. This would require spatial analysis to determine the ideal density and distribution of floodgate stations. Moreover, future studies could examine the role of trust-building mechanisms—such as government transparency in flood response and financial literacy programs—in fostering demand for index-based insurance. Evaluating the effectiveness of policy interventions, such as subsidies and premium adjustments, could also provide valuable insights into strategies for increasing adoption.

Survey questionnaires

Survey on the impacts of Jakarta floods – 2018

A. PROLOGUE				
RESPONDENT PERSONAL INFORMATION				
A00	Respondent ID	:	└─┴─┴─┴─┘ to be filled by the field coordinator	
A01	Name	:	....	
A02	Address	:	....	
A03	City	:	....	└─┴─┘ to be filled by the field coordinator
A04	District (Kecamatan)	:	....	└─┴─┘ to be filled by the field coordinator
A05	Sub-district (Kelurahan)	:	....	└─┴─┘ to be filled by the field coordinator
A06	RT/RW and Post Code	:	└─┴─┘ └─┴─┘ └─┴─┘ └─┴─┘ └─┘	
A07	Landline/mobile phone	:	└─┴─┴─┴─┴─┴─┴─┴─┴─┴─┴─┴─┘	

INTERVIEWER INFORMATION (to be completed by the interviewer)					
A10	Interviewer name and ID	:	……………………………………………		└─┴─┘
A11	Interview date and time (fill in line 2 & 3 if interview is conducted more than once)	:		a. Interview date	b. Start time / Finish time
			1	└─┴─┘ / └─┴─┘ /2017	└─┴─┘ : └─┴─┘ / └─┴─┘ : └─┴─┘
			2	└─┴─┘ / └─┴─┘ /2017	└─┴─┘ : └─┴─┘ / └─┴─┘ : └─┴─┘
			3	└─┴─┘ / └─┴─┘ /2017	└─┴─┘ : └─┴─┘ / └─┴─┘ : └─┴─┘
A12	Interview note	:	1. Only the household head was present during the interview. 2. Both the household head and spouse were present during the interview. 3. Only the spouse of the household head was present during the interview.		└─┘
A13	How strong is the mobile signal at the respondent's home? (Use the mobile phone and SIM card provided by the coordinator)	:	1. One bar 2. Two bars 3. Three bars 4. Four bars 5. Five bars 9. No signal		└─┘
A14	Respondent's GPS coordinates (check GPS)	:	Longitude └─┴─┘ └─┴─┘ └─┴─┘ └─┴─┘ └─┴─┘ Latitude └─┴─┘ └─┴─┘ └─┴─┘ └─┴─┘ └─┴─┘ Elevation └─┴─┘ └─┴─┴─┘ meter ASL (meters above sea level)		
A15	Validation by the Field Coordinator (name and signature)	:			

B. RESPONDENT'S IMMEDIATE FAMILY MEMBER PROFILES

Name of family members (List the names of persons who usually live in the same household and share meals in the same kitchen.)	Relationship to the household head	Sex: 1. Male 2. Female	Age (Number of years)	Marital status	Household members aged 5 or over	
					School participation	Highest level of education attained by each household member
(01)	(02)	(03)	(04)	(05)	(09)	(10)
01. …………………………..						
02. …………………………..						
03. ……………………………						
04. ……………………………						
05. …………………………..						
06. …………………………..						
07. ………………………….						
08. ……………………………						
09. …………………………..						
10. …………………………..						

Family relation code		Marital status code	School participation code	School level code			
1. Head of household 2. Spouse 3. Child 4. Son/Daughter-In-Law 5. Grandchild	6. Parent/In-law 7. Other family member 8. Housemaid 9. Others	1. Unmarried 2. Married 3. Divorced 4. Widow(er)	1. Not yet in school 2. In school 3. Not in school	01. Elementary – 1 02. Elementary – 2 03. Elementary – 3 04. Elementary – 4 05. Elementary – 5	06. Elementary – 6 07. Junior High – 1 08. Junior High – 2 09. Junior High – 3 11. High School – 1	12. High School – 2 13. High School – 3 14. Undergraduate – 1 year 15. Undergraduate – 2 year 16. Undergraduate – 3 year	17. Undergraduate – 4 year 18. Postgraduate (Masters) – 1 year 19. Postgraduate (Masters) – 2 year 20. Postgraduate (Doctorate)

C. SOCIO-ECONOMIC CONDITION (questions for household head)				
House and living condition				
C01	What is the legal status of the house?	:	1. Self-owned → C03 2. Contract 3. Rent 4. Owned by someone else, rent-free → C04 5. Owned by a family member, rent-free → C04 6. Provided by work place → C04 7. Other, please specify .... → C04	└─┘
C02	What is the monthly rent payment for the dwelling? → C05	:	1. IDR └─┴─┘ └─┴─┴─┘ └─┴─┴─┘ /year 2. IDR └─┴─┘ └─┴─┴─┘ └─┴─┴─┘ /month 3. Does not know	└─┘
C03	The status of the land on which the house built is:	:	1. Rights of ownership 2. Rights of building use 3. Rights of usage 4. Other, please specify	└─┘
C04	If you were to rent the house, how much would you need to spend on rent per month/year?	:	1. IDR └─┴─┘ └─┴─┴─┘ └─┴─┴─┘ /year 2. IDR └─┴─┘ └─┴─┴─┘ └─┴─┴─┘ /month 3. Does not know	└─┘
C05	What type of fuel do you use most often for cooking?	:	1. Electricity 2. Gas/LPG 3. Kerosene 4. Firewood 5. Coal 6. Does not cook 7. Other, please specify	└─┘

C. SOCIO-ECONOMIC CONDITION (questions for household head)					
C06	What is your household's primary water source	:	01. Mineral water 02. Tap 03. Well (pump) 04. Well (unprotected/dipper) 05. Spring water 06. Rainwater 07. River 08. Pond 09. Reservoir 10. Other, please specify	a. For drinking	└─┴─┘
				b. For washing dishes	└─┴─┘
C07	What is the household's main water source for sanitation?	:	01. Commercial water 02. Tap water 03. Well (pump) 04. Well (dipper) 05. Spring water 06. Rainwater 07. River 08. Pond 09. Reservoir 10. Other, please specify		└─┴─┘
C08	Does your house have sanitary facilities (bathroom and toilet)?	:	1. Yes 2. No		└─┘

C. SOCIO-ECONOMIC CONDITION (questions for household head)				
C09	Where does this household drain its sewage (from washing, bathing, etc.)?	:	01. Functional sewer 02. Clogged sewer 03. Permanent hole 04. River 05. House yard/garden 06. Pond/lake 07. Puddle 08. Rice field/farm 09. Sea/shore 95. Other, please specify	└─┴─┘
C10	What is the material of the roof of the main living area?	:	1. Concrete 2. Tiles 3. Shingles 4. Zinc 5. Asbestos 6. Palm leaves 7. Other, please specify	└─┘
C11	What is the wall material of the main living area?	:	1. Concrete 2. Wood 3. Bamboo 4. Other, please specify	└─┘
C12	What is the floor material of the main living area?	:	1. Marble/ceramic/granite 2. Tile/terrazzo 3. Cement 4. Wood 5. Dirt 6. Other, please specify	└─┘

C. SOCIO-ECONOMIC CONDITION (questions for household head)				
C13	How many floors does the house have? :	1. 1 floor 2. 2 floors 3. 3 floors 4. More than 3 floors		└─┘
C14	Nearest river to the current house :	a. River name ……………………………		b. Distance from the house └─┴─┴─┴─┘ m
Expenditures				
C15	Household expenditures (IDR) :		Monthly expenditures (last month)	Yearly expenditures (Last 12 months)
		a. Main staple foods (cassava, corn, rice).		
		b. Side dishes, spices, supplementary food, sugar, tea, milk, cooking oil, eggs.		
		c. Additional foods, fruit, snacks, vitamins, etc.		
		D. SUBTOTAL: HOUSEHOLD FOOD EXPENDITURES	(filled in by field coordinator)	
		e. Rent.		
		f. Water, electricity, gas, and kerosene.		
		g. Soap, tooth paste, makeup, shampoo.		
		h. Cigarettes.		
		i. Waste disposal.		
		j. Housemaid services.		
		K. SUBTOTAL: NON-FOOD HOUSEHOLD EXPENDITURES	(filled in by field coordinator)	
		l. Transportation expenses		
		m. Communication expenses (internet, phone, credit)		
		n. Monthly education expense (tuition, stationary, etc.)		

C. SOCIO-ECONOMIC CONDITION (questions for household head)				
		o. Flood-related expenses		
		p. Other daily/weekly expenses (e.g., dues, magazines, etc.)		
		Q. TOTAL MONTHLY EXPENDITURES	(filled in by field coordinator)	
		r. Total household expenses for clothing.		
		s. Expenses for household supplies.		
		t. Health and medication expenses.		
		u. Annual education expenses (books, uniform, school donations, etc.).		
		v. Religious festival expenses (including transportation and gifts).		
		w. Expenses for other goods and services.		
		X. TOTAL ANNUAL FOOD EXPENDITURES		(filled in by field coordinator)
		y. Productive goods and services (investments).		
		z. Land purchases and gifts.		
		aa. Various permits and legalisation (except business permits).		
		ab. Payments for various fines.		
		ac. Savings (set savings, including net savings from rotating savings).		
		ad. Tax (including land and vehicle taxes).		
		ae. Religious and social donations.		
		af. Entertainment and holiday expenses.		
		ag. TOTAL YEARLY NON-FOOD EXPENDITURES		(filled in by field coordinator)
C16	Does any family member have any of the following national health care insurance? (1. Yes; 2. No)	a. BPJS Kesehatan └─┘	b. Jakarta Sehat Card └─┘	c. Jakarta Pintar Card └─┘

C. SOCIO-ECONOMIC CONDITION (questions for household head)							
C17	Other insurances :		a. Life Insurance	b. Health insurance (excluding BPJS)	c. Vehicle insurance	d. House insurance	e. Education insurance
		1. Do any of the family members own any of the following financial schemes? (1. Yes; 2. No)	⎵	⎵	⎵	⎵	⎵
		2. Is the insurance paid by the family member? (1. Yes; 2. No)	⎵	⎵	⎵	⎵	⎵
		3. What is the monthly premium? (IDR)	IDR ………..	IDR ………..	IDR ………..	IDR ……..	IDR ………..
C18	Saving patterns of household head :		1. Save money weekly	2. Save money monthly	3. Save money occasionally	4. Other ………………	
		Do you save your money according to the following patterns? (1. Yes; 2. No)	⎵	⎵	⎵	⎵	
		How much money do you save?	IDR. ……………	IDR. ……………	IDR. ……………	IDR. ………..	
C19	Current assets and liabilities of the household (IDR)	a. Deposit/savings	IDR. …………………………				
		b. Cash	IDR. …………………………				
		c. Money lent	IDR. …………………………				
		d. Personal funds for business and investment	IDR. …………………………				
		e. Loan from a formal financial institution	IDR. …………………………				
		f. Loan from a non-formal financial institution	IDR. …………………………				

D. EXPERIENCE WITH FLOODS				
In this study, a flood event is defined as a situation where streets in the neighborhood are completely inundated, affecting either the entire neighborhood or at least 20 houses in the household's proximity.				
D01	To your knowledge, has your neighborhood experienced a flood in the past ten years (2007–2017)?	:	1. Yes 2. No	└─┘
D02	Do you believe a five-year flood cycle (i.e., occurring every five years) has affected this household?	:	1. Yes 2. No 8. Not sure	└─┘

E. FLOOD CHRONOLOGY IN 2017				
E01	Did the household experience a flood event in 2017?	:	1. Yes 2. No → H01	└─┘

F. IMPACT OF FLOODS IN 2017				
F01	Apart from daily expenses, what additional costs did you face due to flooding? (IDR)	:	01. Food and drinks	IDR └─┴─┘ └─┴─┴─┘ └─┴─┴─┘
			02. Medications	IDR └─┴─┘ └─┴─┴─┘ └─┴─┴─┘
			03. Bathing supplies	IDR └─┴─┘ └─┴─┴─┘ └─┴─┴─┘
			04. Transport fees	IDR └─┴─┘ └─┴─┴─┘ └─┴─┴─┘
			05. General supplies (flashlight, screwdriver, etc.)	IDR └─┴─┘ └─┴─┴─┘ └─┴─┴─┘
			06. Cleaning supplies	IDR └─┴─┘ └─┴─┴─┘ └─┴─┴─┘

F. IMPACT OF FLOODS IN 2017			
		07. Electricity equipment	IDR └─┴─┘ └─┴─┴─┘ └─┴─┴─┘
		08. Temporary accommodation	IDR └─┴─┘ └─┴─┴─┘ └─┴─┴─┘
		09. Vehicle repairs	IDR └─┴─┘ └─┴─┴─┘ └─┴─┴─┘
		10. Furniture repairs	IDR └─┴─┘ └─┴─┴─┘ └─┴─┴─┘
		11. Repair of electrical installations	IDR └─┴─┘ └─┴─┴─┘ └─┴─┴─┘
		12. Sanitation-related repairs	IDR └─┴─┘ └─┴─┴─┘ └─┴─┴─┘
		13 Other houses maintenance repairs	IDR └─┴─┘ └─┴─┴─┘ └─┴─┴─┘
		14. Flood clean-up expenses	IDR └─┴─┘ └─┴─┴─┘ └─┴─┴─┘
		15 Local contributions due to the flood	IDR └─┴─┘ └─┴─┴─┘ └─┴─┴─┘
		95. Other, please specify	IDR └─┴─┘ └─┴─┴─┘ └─┴─┴─┘

G. HOUSEHOLDS MITIGATION AND ADAPTION TO FLOODING				
			a. Action	b. Height
G01	What actions has the household taken to reduce the impact of floods on the current dwelling? (Select all that apply. Mark each selected answer with a "1" and leave blank if not selected.)	:	└─┘ 1. Move house	
			└─┘ 2. Elevate flooring	└─┘ floor (s)
			└─┘ 3. Build a house with two stories or more	
			└─┘ 4. Build a levee (e.g., using sandbags)	└─┴─┴─┘ cm(s)
			└─┘ 5. Move electrical switches and fuse to a higher location	└─┴─┴─┘ cm(s)
			└─┘ 6. Fix the drainage system	
			└─┘ 7. Other, please specify	

G. HOUSEHOLDS MITIGATION AND ADAPTION TO FLOODING				
G02	What did the household do to prepare for the flood? (May choose more than one answer. Mark each answer with a 1, leave blank if not selected.)	:	└─┘ 1. Stock up food supplies └─┘ 2. Stock up clean water supplies └─┘ 3. Place medicines in a storage container └─┘ 4. Leave clothes at relative's/neighbour's house └─┘ 5. Leave valuable documents at relative's/neighbour's house └─┘ 6. Set aside savings for a disaster fund └─┘ 7. Buy insurance └─┘ 9. Other, please specify ……………….	
G03	What actions is this household going to take regarding the current dwelling in order to reduce the impact of the flood(s)? (May choose more than one answer. Mark each answer with a 1, leave blank if not selected.)	:	a. Action	b. Height
			└─┘ 1. Move house	
			└─┘ 2. Elevate house flooring	└─┘ floor(s)
			└─┘ 3. Build a house with two stories or more	
			└─┘ 4. Build a levee (e.g., using sandbags)	└─┴─┴─┘ cm(s)
			└─┘ 5. Move electrical switches and fuse to a higher location	└─┴─┴─┘ cm(s)
			└─┘ 6. Fix the drainage system	
			└─┘ 7. Other, please specify ……………….	
G04	What do local residents do to reduce the impact of floods on their dwellings? (Rank from most to least common)	:	a. Move house b. Elevate house flooring c. Build a house with two stories or more d. Build a levee (e.g., using sandbags) e. Move electrical switches and fuse to a higher location f. Fix the drainage system g. Other, please specify ……………….	1. └─┘ 2. └─┘ 3. └─┘ 4. └─┘ 5. └─┘ 6. └─┘ 7. └─┘

G. HOUSEHOLDS MITIGATION AND ADAPTION TO FLOODING			
G05	[If G03 = 1] What has the household been doing to relocate?	:	└─┘ 1. Searching/Surveying other locations not affected by floods └─┘ 2. Saving money to buy another house └─┘ 3. Contacting banks/financial institution for the possibility of getting a home loan └─┘ 4. Putting the house up for sale to buy another house └─┘ 5. Other, please specify ………………

H. THE WORST FLOOD EXPERIENCE (2012–2016)			
H01	During the period 2012–2016, when did the household experience the worst flood?	:	a. Year └─┴─┴─┴─┘ b. Month └─┴─┘
H02	In monetary terms, what were the additional expenses due to this flood?	:	IDR ……………………………………………
Go to K01			

I. IMPACTS OUTSIDE THE FLOODING ZONE				
I01	In the past six years (2012–2017), when did the worst flooding occur in Jakarta?	:	Year └─┴─┴─┴─┘ Month └─┴─┘	
The following questions are related to general flooding issues				
I02	What actions is this household planning to take on the current dwelling to reduce the impact of flooding? (You may choose more than one answer. Mark	:	a. Action	b. Height
			└─┘ 1. Move house	
			└─┘ 2. Elevate flooring	└─┘ floors
			└─┘ 3. Build a house with two stories or more	
			└─┘ 4. Build a levee (e.g., using sandbags)	└─┴─┴─┘ cm(s)

I. IMPACTS OUTSIDE THE FLOODING ZONE						
	each selected answer with a "1" and leave empty if not selected).	└─┘ 5. Move electrical switches and fuse to a higher location			└─┴─┴─┘ cm(s)	
		└─┘ 6. Fix the drainage system				
		└─┘ 7. Other, please specify				
I03	[If I02 = 1] What has the household been doing to move house? :	└─┘ 1. Searching/Surveying other locations not affected by floods └─┘ 2. Saving money to buy another house └─┘ 3. Contacting banks/financial institutions about the possibility of getting a home loan └─┘ 4. Selling current house on the market to buy another house └─┘ 5. Other, please specify				
I04	Financial schemes to minimise the impact of floods: :		a. Education deposit	b. Health deposit	c. Flood insurance	d. Nothing
		1. Does the household have […] to minimise the impact? (1. Yes, 2. No)	└─┘	└─┘	└─┘	└─┘
		2. Amount per month	IDR…………	IDR…………	IDR …………	

K. INDEX-BASED FLOOD INSURANCE			
K01	Sluice/Floodgate	In 2007, the water level at the sluice gate [NAME], the closest flood gate to your house, exceeded the critical level of [CRITICAL LEVEL] (the maximum water level for "*Siaga 1*"), causing a flood in Jakarta with an average height of 40 cm (see the flood map). This information was obtained from the official in charge of the sluice/floodgate [NAME]. ---	└─┘

Note: The interviewer must choose one of the following pairs:
[NAME] is the nearest sluice/floodgate, and [CRITICAL LEVEL] is the water level height for the critical level of Siaga 1.

[NAME]	[CRITICAL LEVEL]
Katulampa Dam	200 cm
Depok Post	350 cm
Manggarai Sluice	950 cm
Karet Sluice	600 cm
Krukut Hulu Post	300 cm
Sunter Hulu Post	250 cm
Cipinang Hulu Post	250 cm
Pulogadung Sluice	770 cm
Cakung Drain Sluice	390 cm
Angku Hulu Post	300 cm
Pesanggrahan Post	350 cm
Cengkareng Drain Sluice	310 cm

K02	Index-based flood insurance	The DKI Jakarta provincial government is considering the implementation of a flood mitigation program based on its current available budget (i.e., a business-as-usual case). Consequently, the probability of flooding in Jakarta remains approximately the same as what you have experienced in recent years. The DKI Jakarta provincial government plans to introduce an index-based flood insurance program for all residents of Jakarta who are interested. This insurance scheme will only be implemented if a large number of people are willing to participate. **How the index-based flood insurance works:** If the current water level registered at Floodgate Station [NAME] reaches a height of [CRITICAL LEVEL] or more, all insured households will receive flood compensation/assistance of IDR 10 million. Taking into account your current household income and expenditure levels, are you willing to participate in this index-based flood insurance program by paying a [WTP] premium per month? 1. Yes → K04 2. No **Notes:** [WTP] is randomly selected by Survey Solutions from the following list: 1. IDR 10,000 2. IDR 50,000 3. IDR 100,000 4. IDR 200,000 5. IDR 500,000	└─┘

K03	Are you willing to participate in the flood insurance program by paying a monthly premium of IDR [WTP]/2?	1. Yes → K06 2. No → K05	└─┘
K04	Are you willing to participate in the flood insurance program by paying a monthly premium of IDR 2x[WTP]?	1. Yes → K06 2. No → K06	└─┘
K05	What are the reasons you are not willing to participate in the insurance program?	1. Does not have the money to pay the monthly premium 2. Does not trust or believe that the DKI Jakarta provincial government can manage the insurance funds transparently 3. Have never had any insurance; prefer to use money for personal savings 4. Does not trust the insurance system 5. Other, please specify	└─┘
K06	In your opinion, among the following institutions, which one do you trust the most to manage the insurance funds?	1. Flood insurance organised by the DKI Jakarta provincial government 2. Flood insurance organised by a private insurance company 3. Flood insurance organised by NGOs that have been helping to manage flood situations 4. Other, please specify	└─┘
K07	If this index-based flood insurance is implemented, what factors do you think are the main contributors to the success of the program? (Please rank the four main contributors from most important to least important. Interviewer, please read the options.)	1. Insurance institution 2. Amount of compensation received 3. Amount of premium paid 4. Compensation procedures 5. Length of time between filing a claim and disbursement of compensation funds 6. Other, please specify	1. └─┘ 2. └─┘ 3. └─┘ 4. └─┘

L. RISK ATTITUDES	
L01	**Game I:** We are going to play a game. The game is part of our research to understand how people make decisions and their willingness to take risks in general. **There are no right or wrong decisions. You should just make the decision that feels right to you. Here is what we are going to do:** Please note that **no real money** is involved in this game. Suppose I give you IDR 20,000, and you have six options to invest this amount. Each of these investments has a 50% chance of yielding a return as shown in Case A, and a 50% chance of yielding a return as shown in Case B. Please show the CARD. For example: If you pick Investment Option 2, there is a 50% chance that you will receive a return of IDR 16,000 and a 50% chance that you will receive a return of IDR 25,000. Or, if you pick Option 5, there is a 50% chance that you will receive a return of IDR 4,000 and a 50% chance that you will receive a return of IDR 40,000. Knowing the possible returns of each investment option, which investment option would you choose? Investment number: └─┘ Next, we are going to calculate your winnings. Whichever investment option you choose, your winnings will be determined by us playing "which-hand-is-it in." Here, I have two marbles: a blue one and a yellow one. I am going to put them behind my back, shake them, and then place one in my left hand and the other in my right hand. After that, I will bring both of my hands forward. So now, I have one marble concealed in each hand.

Before we go any further, which do you think is more likely: picking the hand with the blue marble or the hand with the yellow marble?

1. Blue marble
2. Yellow marble

If you pick the hand containing the blue marble, you will be in Case A. If you pick the hand containing the yellow marble, you will be in Case B.

Now, pick one of my hands. The colour of the marble in the hand you pick is:

1. BLUE
2. YELLOW

And so, your return for the investment option you picked is IDR: ………………………

INVESTMENT OPTIONS :

	Case A (Blue Marble)	Case B (Yellow Marble)
1	IDR 20,000	IDR 20,000
2	IDR16,000	IDR 25,000
3	IDR 12,000	IDR 30,000
4	IDR 8,000	IDR 35,000
5	IDR 4,000	IDR 40,000
6	IDR 0	IDR 50,000

└─┘

L02	**Game II:** We are going to play the same game as the previous one. However, this time, we will actually give you the money based on your investment option. **Please note that there are no right or wrong decisions. You just make the decision that feels right to you. This is what we are going to do.** I will give you IDR 20,000, but you need to invest this IDR 20,000 into one of the six investment options. Each of these investments has a 50% probability of yielding a return as much as stated in Case A, and a 50% probability of yielding a return as much as stated in Case B. Please show the CARD. For example: If you pick Investment Option 2, there is a 50% probability that you will receive a return of IDR 16,000, and a 50% probability that you will receive a return of IDR 25,000. Or, if you pick Option 5, there is a 50% probability that you will receive a return of IDR 4,000, and a 50% probability that you will receive a return of IDR 40,000. Knowing the possibilities of returns for each investment option, which investment option would you choose? Investment number: └─┘ RESPONDENT'S SIGNATURE ______________________ Next, we are going to calculate your winnings. Whichever investment option you choose, your winnings will be determined by us playing "which-hand-is-it in." Here, I have two marbles: a blue one and a yellow one. I am going to put them behind my back, shake them, then place one of them in my left hand and the other in my right hand. After that, I will bring both of my hands forward. So, now I have one marble concealed in each hand. You must pick one of my hands. If you pick the hand containing the blue marble, you will be in Case A. If you pick the hand containing the yellow marble, you will be in Case B.

	Before we go any further, which do you think is more likely: that you will pick the hand with the blue marble or the hand with the yellow marble? 1. Blue marble 2. Yellow marble Now, pick one of my hands. The colour of the marble in the hand you pick is: 1. Blue 2. Yellow └─┘ RESPONDENT'S SIGNATURE ______________________________ And so, your return for the investment you picked is IDR: ………………………. Here is the money.

Household survey on the impact of floods in Lima – 2023

A. PROLOGUE				
RESPONDENT PERSONAL INFORMATION				
A00	Respondent ID	:	└─┴─┴─┴─┘ to be filled by the field coordinator	
A01	Name	:	....	
A02	Address	:	....	
A03	City	:	....	└─┴─┘ to be filled by the field coordinator
A04	District	:	....	└─┴─┘ to be filled by the field coordinator
A05	Neighbourhood unit	:	....	└─┴─┘ to be filled by the field coordinator
A06	Community unit	:	....	└─┴─┘ to be filled by the field coordinator
A07	Post code	:	└─┴─┴─┴─┴─┘	
A08	Landline/mobile phone	:	└─┴─┴─┴─┴─┴─┴─┴─┴─┴─┴─┴─┘	

<table>
<tr><th colspan="5">INTERVIEWER INFORMATION (to be completed by the interviewer)</th></tr>
<tr><td>A09</td><td>Interviewer name and ID</td><td>:</td><td colspan="2">……………………………………………………………………………… └─┴─┘</td></tr>
<tr><td rowspan="4">A10</td><td rowspan="4">Interview date and time
(fill in line 2 & 3 if the interview is conducted more than once)</td><td rowspan="4">:</td><td></td><td>a. Interview date | b. Start time / Finish time</td></tr>
<tr><td>1</td><td>└─┴─┘ / └─┴─┘ /2023 | └─┴─┘ : └─┴─┘ / └─┴─┘ : └─┴─┘</td></tr>
<tr><td>2</td><td>└─┴─┘ / └─┴─┘ /2023 | └─┴─┘ : └─┴─┘ / └─┴─┘ : └─┴─┘</td></tr>
<tr><td>3</td><td>└─┴─┘ / └─┴─┘ /2023 | └─┴─┘ : └─┴─┘ / └─┴─┘ : └─┴─┘</td></tr>
<tr><td>A11</td><td>Interview note</td><td>:</td><td colspan="2">1. Only the household head was present during the interview.
2. Both the household head and spouse were present during the interview. └─┘
3. Only the spouse of the household head was present during the interview.</td></tr>
<tr><td>A12</td><td>How strong is the mobile signal at the respondent's home?
(Use the mobile phone and SIM card provided by the coordinator)</td><td>:</td><td colspan="2">1. One bar
2. Two bars
3. Three bars
4. Four bars
5. Five bars
9. No signal
└─┘</td></tr>
<tr><td>A13</td><td>Respondent's GPS coordinates
(check GPS)</td><td>:</td><td colspan="2">Longitude └─┴─┘ └─┴─┘ └─┴─┘ └─┴─┘ └─┴─┘
Latitude └─┴─┘ └─┴─┘ └─┴─┘ └─┴─┘ └─┴─┘
Elevation └─┴─┘ └─┴─┴─┘ meter ASL (meters above sea level)</td></tr>
<tr><td>A14</td><td>Validation by the Field Coordinator
(name and signature)</td><td>:</td><td colspan="2"></td></tr>
</table>

B. RESPONDENT'S IMMEDIATE FAMILY MEMBER PROFILES									
Name of family members (List the names of persons who usually live in the same household and share meals in the same kitchen.)	Relationship to the household head	Sex: 1. Male 2. Female	Age (Number of years)	Marital status	Household members aged 10 and over			Household members aged 5 or over	
					Did (s)he work **last week**? 1. Yes 2. No	Does (s)he work or own a business that is temporarily not operating **within this week**? 1. Yes 2. No	Was (s)he looking for a job or planning to start a business **last week**? 1. Yes 2. No	School participation	Highest level of education attained by each household member
(01)	(02)	(03)	(04)	(05)	(06)	(07)	(08)	(09)	(10)
01. ………………	└─┘	└─┘	└─┴─┘	└─┘	└─┘	└─┘	└─┘	└─┘	└─┘
02. ………………	└─┘	└─┘	└─┴─┘	└─┘	└─┘	└─┘	└─┘	└─┘	└─┘
03. ………………	└─┘	└─┘	└─┴─┘	└─┘	└─┘	└─┘	└─┘	└─┘	└─┘
04. ………………	└─┘	└─┘	└─┴─┘	└─┘	└─┘	└─┘	└─┘	└─┘	└─┘
05. ………………	└─┘	└─┘	└─┴─┘	└─┘	└─┘	└─┘	└─┘	└─┘	└─┘
06. ………………	└─┘	└─┘	└─┴─┘	└─┘	└─┘	└─┘	└─┘	└─┘	└─┘
07. ………………	└─┘	└─┘	└─┴─┘	└─┘	└─┘	└─┘	└─┘	└─┘	└─┘

C. EMPLOYMENT STATUS (*Only for working household members, if B06=1 or B07=1*)						
ID Number of Household Member (*See section C, column (1)*)	Number of working days (main occupation) **last week** = …… days	Number of hours worked (main occupation) last week = …… hours	What is the (main) type of business or field of work of the workplace (name) **within the last week**?	Position in main occupation **within the last week**:	Does the occupation provide in-kind benefit/salary? 1. Yes 2. No	Does (s)he have secondary occupation/secondary business 1. Yes 2. No
(01)	(02)	(03)	(04)	(05)	(06)	(08)
└─┴─┘	└─┴─┘	└─┴─┘	└─┴─┘	└─┘	└─┘	└─┘
└─┴─┘	└─┴─┘	└─┴─┘	└─┴─┘	└─┘	└─┘	└─┘
└─┴─┘	└─┴─┘	└─┴─┘	└─┴─┘	└─┘	└─┘	└─┘
└─┴─┘	└─┴─┘	└─┴─┘	└─┴─┘	└─┘	└─┘	└─┘
└─┴─┘	└─┴─┘	└─┴─┘	└─┴─┘	└─┘	└─┘	└─┘

Family relation code	Marital status code	School participation code	School level code		Field of work code		Status/position code
1. Head of household 2. Spouse 3. Child 4. Son/Daughter-In-Law 5. Grandchild 6. Parent/In-law 7. Other family member 8. Housemaid 9. Others	1. Unmarried 2. Married 3. Divorced 4. Widow(er)	1. Not yet in school 2. In school 3. Not in school	01. Elementary – 1 02. Elementary – 2 03. Elementary – 3 04. Elementary – 4 05. Elementary – 5 06. Elementary – 6 07. Junior High – 1 08. Junior High – 2 09. Junior High – 3 11. High School – 1 12. High School – 2 13. High School – 3 14. Undergraduate – 1 year	15. Undergraduate – 2 years 16. Undergraduate – 3 years 17. Undergraduate – 4 years 18. Postgraduate (Masters)– 1 year 19. Postgraduate (Masters) – 2 years 20. Postgraduate (Doctorate)	01. Agriculture 02. Horticulture 03. Plantation 04. Fishery 05. Farming 06. Forestry & other agriculture 07 Mining & excavation 08. Processing industry 09. Gas & electricity 10. Construction/Building 11. Trade	12. Food and Hospitality 13. Transportation and Storage 14. Information and Communication 15. Accounting and Insurance 16. Education services 17. Health services 18. Social, governmental & personal services 19. Other	1. Business owner 2. Business owner, assisted by part-time/volunteer workers 3. Business owner, assisted by full time/paid workers 4. Company worker 5. Freelancer 6. Family business or unpaid worker

D. SOCIO-ECONOMIC CONDITION (questions for Household Head)				
House and living condition				
D01	What is the status of the house?	:	1. Self-owned (with property title) →D03 2. Self-owned (without property title) →D05 3. Contract 4. Rent 5. Owned by someone else, rent-free → D05 6. Owned by a family member, rent-rent → D05 7. Provided by work place → D05 8. Other …………………………………… → D05	└┘
D02	What is the monthly rent payment for the dwelling? → D06	:	1. PEN └┴┴┘ └┴┴┴┘ └┴┴┴┘ /year 2. PEN └┴┴┘ └┴┴┴┘ └┴┴┴┘ /month 3. Does not know	└┘
D03	In what year was the property title granted to you?	:	in year └┴┴┴┘	
D04	What type of property right does the household head have over the land where the house was built?	:	1. Full ownership 2. Usufruct 3. Bare ownership 4. Other, please specify	└┘
D05	If you were to rent the house, how much would you have to pay per month/year in rent?	:	1. PEN └┴┴┘ └┴┴┴┘ └┴┴┴┘ /year 2. PEN └┴┴┘ └┴┴┴┘ └┴┴┴┘ /month 3. Not sure	└┘

D. SOCIO-ECONOMIC CONDITION (questions for Household Head)					
D06	What is the main source you use to light your house?	:	1. Electricity from Luz del Sur (Private Electricity Company) 2. Another Electricity Company 3. Lantern 4. Lamp/torch 5. Generator 6. Other, please specify		└─┘
D07	What fuel do you use most frequently for cooking?	:	1. Electricity 2. Gas/LPG (gas cylinder) 3. Gas/LPG (pipeline) 4. Kerosene 5. Coal 6. Wood 7. Dung 8. Does not cook 9. Other, please specify		└─┘
D08	What is the main source of water in the household?	:	1. Branded bottled water 2. Tap water connection 3. Well (closed) 4. Well (open) 5. Spring water 6. Rainwater 7. River water 8. Stream water 9. Creek water 10. Pond 11. Reservoir 12. Water tanker 13. Other, please specify	a. For drinking b. For washing dishes	└─┴─┘ └─┴─┘

D. SOCIO-ECONOMIC CONDITION (questions for Household Head)				
D09	Does your house have sanitary facilities (bathroom and toilet)?	:	1. Yes 2. No	└─┘
D10	Where is the wastewater (water used for washing/bathing, etc.) from this household disposed of?	:	1. Sewer 2. Permanent hole/pit 3. River 4. Household yard/garden 5. Pond/lake 6. Puddle 95. Other, please specify	└─┴─┘
D11	What is the material of the floor in the main living room?	:	1. Parquet/polished wood 2. Asphalt sheets/vinyl/similar 3. Tiles/terrazzo/ceramics/similar 4. Wood (pine, screw pine, etc.) 5. Cement 6. Earth 7. Other, please specify	└─┘
D12	How many floors does the house have?	:	1. 1 floor 2. 2 floors 3. 3 floors 4. More than 3 floors	└─┘

D. SOCIO-ECONOMIC CONDITION (questions for Household Head)			
Expenditures			
D13	Household expenditures (PEN) :	a. Approximately how much did you and your family spend on food during the LAST WEEK? Include expenses for drinking water, spices, herbs, cooking oil, and other cooking fuels (firewood, LPG gas, kerosene).	PEN └─┴─┴─┘
		b. How much do you and your family spend on cigarettes per day?	PEN └─┴─┴─┘
		c. How much did you and your family spend on alcoholic beverages during the LAST WEEK?	PEN └─┴─┴─┘
		d. How much did you and your family spend on mobile phones during the LAST WEEK?	PEN └─┴─┴─┘
		e. How much did you and your family spend on electricity in the LAST MONTH?	PEN └─┴─┴─┘
		f. Approximately how much did you and your family spend on transportation in the LAST MONTH? Include public transportation expenses, and if any family member owns a car/motorcycle, include expenses for fuel, repairs, and taxes.	PEN └─┴─┴─┘
		g. Approximately how much did you and your family spend on health services in the LAST MONTH? Include expenses for medications and visits to doctors and nurses.	PEN └─┴─┴─┘
		h. How much did you and your family spend on education during the LAST MONTH? Include tuition and school activity expenses.	PEN └─┴─┴─┘
		i. How much did you and your family spend on entertainment during the LAST MONTH? Include expenses for dining out, drinking with friends/family, and other recreational activities.	PEN └─┴─┴─┘
		j. How much did you and your family spend on soap, detergent, and other cleaning products in the LAST MONTH?	PEN └─┴─┴─┘

D. SOCIO-ECONOMIC CONDITION (questions for Household Head)			
D14		k. Approximately how much did you and your family spend on loan payments in the LAST MONTH?	PEN └─┴─┴─┘
		l. Approximately how much did you and your family spend on house maintenance in the LAST 6 MONTHS?	PEN └─┴─┴─┘
		m. Approximately how much did you and your family spend on clothing (including shoes and garments) in the LAST YEAR?	PEN └─┴─┴─┘
		n. Approximately how much did you and your household spend on ceremonial activities in the LAST YEAR? Include expenses for religious and cultural ceremonies, parties, and weddings.	PEN └─┴─┴─┘
		o. Approximately how much did you and your family spend on renting your house in the LAST YEAR?	PEN └─┴─┴─┘
		p. Approximately how much did you and your family spend in total in the LAST MONTH?	PEN └─┴─┴─┘
D14	Does any member of this household have any of the following public health insurance? :	1. SIS (Comprehensive Health Insurance) 2. EsSalud (Social Health Insurance)	└─┘

D15	Other insurances :	Does any family member have any other types of insurance? (1. Yes; 2. No)	a. Life insurance	b. Private health insurance or EPS	c. Vehicle insurance	d. House insurance	e. Education insurance
			└─┘	└─┘	└─┘	└─┘	└─┘

E. EXPERIENCE WITH FLOODS			
In this study, floods are sudden events caused by heavy rainfall that occur in the highland regions between January and April. They bring water, mud, and debris, flooding the streets around the entire neighborhood or at least 20 houses around your home.			
E01	In your opinion, has the neighborhood where you live experienced floods between the years 2015 and 2023? :	1. Yes 2. No	└─┘
E02	In what year did the flood occur? :	1. 2015 2. 2017 3. 2023	└─┘ └─┘ └─┘

F. CHRONOLOGY OF THE LAST FLOOD			
F01	When was the last time this neighbourhood experienced a flood? :	└─┘ 1. Before 2012 └─┘ 2. 2012 └─┘ 3. 2013-2014 └─┘ 4. 2015 └─┘ 5. 2016 └─┘ 6. 2017 └─┘ 7. 2018-2022 └─┘ 8. 2023	
F02	When the flood occurred, what additional expenses were added to the household's daily expenses? (PEN) :	01. Food and drinks	PEN └─┴─┘ └─┴─┴─┘
		02. Medicines	PEN └─┴─┘ └─┴─┴─┘
		03. Toiletries	PEN └─┴─┘ └─┴─┴─┘

F. CHRONOLOGY OF THE LAST FLOOD			
		04. Mobility/transportation costs	PEN └─┴─┘ └─┴─┴─┘
		05. General tools (flashlight, screwdriver, etc.)	PEN └─┴─┘ └─┴─┴─┘
		06. Cleaning supplies	PEN └─┴─┘ └─┴─┴─┘
		07. Vehicle repair	PEN └─┴─┘ └─┴─┴─┘
		10. Furniture repairs	PEN └─┴─┘ └─┴─┴─┘
		11. Repair of electrical installations	PEN └─┴─┘ └─┴─┴─┘
		95. Other, please specify	PEN └─┴─┘ └─┴─┴─┘
		99. None	

G. MITIGATION AND ADAPTATION OF HOUSEHOLDS TO FLOODING				
			a. Action	b. Height
G01	What actions has the household taken to reduce the impact of flooding on the current home? (You can choose more than one option. Leave blank if the option was not selected.)	:	└─┘ 1. Move house	
			└─┘ 2. Elevate flooring	└─┘ floor (s)
			└─┘ 3. Build a house with two stories or more	
			└─┘ 4. Build a levee (e.g., using sandbags)	└─┴─┴─┘ cm(s)
			└─┘ 5. Fix the drainage system	
			└─┘ 6. Other, please specify ……………….	
			└─┘ 7. None	
G02	What did the household do to prepare for the floods? (You can choose more than one option. Leave blank if the option was not selected.)	:	└─┘ 1. Stocked up on food └─┘ 2. Stocked up on drinking water └─┘ 3. Stored medicines └─┘ 4. Left clothes at friends'/relatives' house └─┘ 5. Left valuable documents at friends'/relatives' house	

G. MITIGATION AND ADAPTATION OF HOUSEHOLDS TO FLOODING		
		└─┘ 6. Saved an emergency fund for disasters └─┘ 7. Purchased insurance └─┘ 8. Other, please specify ………………. └─┘ 9. Did not make any preparations

H. IMPACTS OUTSIDE THE FLOODING ZONE				
H01	If the house floods, what actions would the household take to reduce the impact of the flooding? (You may choose more than one option. Leave blank if the option is not selected.)	:	a. Action	b. Height
			└─┘ 1. Would move to a new house	
			└─┘ 2. Would raise the floor level of the house	└─┘ floors
			└─┘ 3. Would build a two-story house or higher	└─┴─┴─┘ cm(s)
			└─┘ 4. Would build a levee (e.g., using sand/cement bags)	└─┴─┴─┘ cm(s)
			└─┘ 5. Would fix the drainage system	
			└─┘ 6. Other, please specify ……………….	
			└─┘ 7. None	
H02	What did the household do to prepare for the floods? (You can choose more than one option. Leave blank if the option was not selected.)	:	└─┘ 1. Stocked up on food └─┘ 2. Stocked up on drinking water └─┘ 3. Stored medicines └─┘ 4. Left clothes at friends'/relatives' house └─┘ 5. Left valuable documents at friends'/relatives' house └─┘ 6. Saved an emergency fund for disasters └─┘ 7. Purchased insurance └─┘ 8. Other, please specify ………………. └─┘ 9. Did not make any preparations	

I. RISK ATTITUDES	
I01	Now we are going to play a game. This game is part of our research to understand how people make decisions and take risks in general. **There are no right or wrong answers—simply choose the option that seems best to you.** Please note that **real money** is involved in this game. The description of the game is as follows: Suppose I give you PEN 2.00, and you have six investment options. Please show the CARD. Here are the six investment options: Option 1, Option 2, Option 3, Option 4, Option 5, and Option 6. Each investment option has a 50% probability of receiving an amount of money shown in the BLUE column and a 50% probability of receiving a payment shown in the YELLOW column. Let me give you an example of this game. Imagine that you invest the PEN 2.00 in Option 3. This means there is a 50% chance of receiving a payment of PEN 1.20 and a 50% chance of receiving a payment of PEN 3.00. In other words, in the first scenario, you would lose PEN 0.80, and in the second, you would gain PEN 1.00. In a second example, if you invest the PEN 2.00 in Option 5, there is a chance you will receive a payment of PEN 0.40 and another chance of receiving a payment of PEN 4.00. In the first case, you would lose PEN 1.60, while in the second, you would gain PEN 2.00. After this explanation, which investment option will you choose? Investment number: └─┘ Next, we are going to calculate your winnings. Whichever investment option you choose, your winnings will be determined by playing "Which-hand-is-it?" Here I have two marbles, one BLUE and one YELLOW. I will put them behind my back and shake them, then place one in my left hand and the other in my right hand. I will then bring both of my hands forward, so now I have one marble concealed in each hand.

Before we go any further, which do you think is more likely: that you pick the hand with the blue marble or the hand with the yellow marble?

1. Blue marble
2. Yellow marble

If you pick the hand containing the blue marble, you will be in Case A. If you pick the hand containing the yellow marble, you will be in Case B.

Now, pick one of my hands. The colour of the marble in the hand you pick is:

1. BLUE
2. YELLOW └─┘

And so, your return for the investment option you chose is PEN ……………………….. Here is the money.

RESPONDENT'S SIGNATURE ____

INVESTMENT OPTIONS :

	Case A (Blue Marble)	Case B (Yellow Marble)
1	PEN 5.00	PEN 5.00
2	PEN 4.00	PEN 6.50
3	PEN 3.00	PEN 8.00
4	PEN 2.00	PEN 9.00
5	PEN 1.00	PEN 10.50
6	PEN 0.00	PEN 12.50

└─┘

J. TIME PREFERENCE	
J01	Game: This game is part of our research to understand how people make decisions and their level of patience in general. There are no right or wrong decisions; you are simply asked to make the choice that seems best to you. Here, the money is **<u>not real</u>**. Imagine that you will hear two different payment options that someone wants to offer you. One payment would be received in the near future (for example, TOMORROW), and the other in the distant future (for example, in 3 months) but for a higher amount. Finally, that person asks you: *Which payment would you like to receive?* You can only choose one of the payments. Now, please answer the following questions: • Imagine someone offers you PEN 5.00 tomorrow or PEN 5.30 in 3 months. Which one do you choose? • Second scenario: Imagine someone offers you PEN 5.00 tomorrow or PEN 6.00 in 3 months. Which one do you choose? • Third scenario: Imagine someone offers you PEN 5.00 tomorrow or PEN 7.50 in 3 months. Which one do you choose? • Fourth scenario: Imagine someone offers you PEN 5.00 tomorrow or PEN 10.00 in 3 months. Which one do you choose? --- Note: The order in which each choice appears on your screen is randomly selected by Survey Solutions.

Time preference

	Option 1	Option 2	Answer
Choice 1	PEN 5 soles tomorrow	PEN 5.30 soles in 3 months	⎵
	PEN 5 soles tomorrow	PEN 6.00 soles in 3 months	⎵
	PEN 5 soles tomorrow	PEN 7.50 soles in 3 months	⎵
	PEN 5 soles in 1 month	PEN 10.00 soles in 3 months	⎵
Choice 2	PEN 5 soles in 1 month	PEN 5.30 soles in 4 months	⎵
	PEN 5 soles in 1 month	PEN 6.00 soles in 4 months	⎵
	PEN 5 soles in 1 month	PEN 7.50 soles in 4 months	⎵
	PEN 5 soles in 1 month	PEN 10.00 soles in 4 months	⎵
Choice 3	PEN 5 soles in 1 year	PEN 5.30 soles in 1 year 3 months	⎵
	PEN 5 soles in 1 year	PEN 6.00 soles in 1 year 3 months	⎵
	PEN 5 soles in 1 year	7.50 soles in 1 year 3 months	⎵
	PEN 5 soles in 1 year	PEN 10.00 soles in 1 year 3 months	⎵

K. INDEX-BASED FLOOD INSURANCE			
K01	Index-based flood insurance	Imagine that the local government of Lurigancho-Chosica is planning to launch a new type of insurance to cover household's property losses caused by flooding. This insurance is known as index-based flood insurance, which is different from traditional insurance. In traditional insurance—such as accident insurance—the policyholder must lodge a claim, undergo a damage assessment, and then receive a payout. If the claim is accepted, the policyholder receives compensation; if the claim is denied, they receive nothing. In contrast, index-based insurance provides compensation automatically after the disaster occurs. This insurance is based on a predefined parameter that indicates the rainfall threshold at which severe flooding and property damage occur in a given area. This threshold is monitored by a weather station of Lurigancho-Chosica. For example, if the weather station records rainfall in a specific area that exceeds the known threshold for causing devastating floods, households that purchased the insurance would receive compensation. However, if rainfall does not exceed that threshold, insured households would not receive any payout. Thus, the weather station determines whether rainfall has crossed the threshold known to cause floods and property damage. In the case of the Riverbed [NAME], the rainfall threshold for flooding and damage is [ACCUMULATED RAINFALL] millimetres (mm). In 2017, rainfall exceeded this threshold, and households suffered the impact of mudslides (*huaycos*). Now, if the weather station reports that rainfall in Riverbed [NAME] crosses the height of [ACCUMULATED RAINFALL] in three days in a row, households	⎵

<table>
<tr><td></td><td></td><td>that purchased index-based insurance would receive a compensation of PEN 10,000.

<table>
<tr><th>[NAME]</th><th>[ACCUMULATED RAINFALL]</th></tr>
<tr><td>Nicolás de Piérola</td><td rowspan="3">36 mm</td></tr>
<tr><td>Sierra Limeña</td></tr>
<tr><td>San Antonio Pedregal</td></tr>
</table>

Taking into account your current household income and expenditures levels, are you willing to participate in this index-based flood insurance program by paying [WTP] premium per month?

1. Yes → K03
2. No

Note:
[WTP] is randomly selected by Survey Solutions from the following list:
1. PEN 10.00
2. PEN 20.00
3. PEN 50.00
4. PEN 200.00
5. PEN 500.00</td><td></td></tr>
<tr><td>K02</td><td>Are you willing to participate in the flood insurance program by paying a monthly premium of PEN [WTP]/2?</td><td>1. Yes → K05
2. No → K04</td><td>└─┘</td></tr>
<tr><td>K03</td><td>Are you willing to participate in the flood insurance program by paying a monthly premium of PEN 2x[WTP]?</td><td>1. Yes → K05
2. No → K05</td><td>└─┘</td></tr>
</table>

K04	What are the reasons you are not willing to participate in the insurance program?	1. Do not have the money to pay the monthly premium 2. Do not trust that the local government of Lurigancho-Chosica can manage the insurance funds transparently 3. Have never had any insurance; it is better to use money as personal savings 4. Do not trust the insurance system 5. Other, please specify.................	└─┘
K05	In your opinion, among the following institutions, which one do you trust the most to manage the insurance funds?	1. Flood insurance organized by the local government of Lurigancho-Chosica 2. Flood insurance organized by a private insurance company 3. Flood insurance organized by NGOs that have been helping to manage floods 4. Other, please specify	└─┘
K06	If this index-based flood insurance is implemented, what factors do you think are the main contributors to the success of the program? (Please rank the four main contributors from most important to least important. Interviewer, please read the options.)	1. Insurance institution 2. Amount of compensation received 3. Amount of premium paid 4. Compensation procedures 5. Length of time between claiming and disbursement of compensation funds 6. Other, please specify	1. └─┘ 2. └─┘ 3. └─┘ 4. └─┘

L. CURRENT LOCAL GOVERNMENT PROGRAM			
L01	a. In your opinion, do you agree with the current local government's program for infrastructure development, education, and health?	1. Strongly agree 2. Agree 3. Disagree 4. Strongly disagree	└─┘
L02	b. Can the local government be trusted to effectively manage floods?	1. Strongly agree 2. Agree 3. Disagree 4. Strongly disagree	└─┘