Household level effects of flooding: Evidence from Thailand

Zhuldyz Ashikbayeva Marei Fürstenberg Timo Kapelari Albert Pierres
Stephan Thies

June 9, 2019

Abstract

This thesis studies the impacts of flooding on income and expenditures of rural households in northeast Thailand. It explores and compares shock coping strategies and identifies household level differences in flood resilience. Drawing on unique household panel data collected between 2007 and 2016, we exploit random spatio-temporal variation in flood intensities on the village level to identify the causal impacts of flooding on households. Two objective measures for flood intensities are derived from satellite data and employed in the analysis. Both proposed measures rely on the percentage area inundated in the surrounding of a village, but the second measure is standardized and expressed in comparison to the median village level flood exposure. We find that household incomes are negatively affected by floods. However, our results suggest that rather than absolute levels of flooding, deviations from median flood exposure are driving negative effects on households. This indicates a certain degree of adaptation to floods. Household expenditures for health and especially food rise in the aftermath of flooding. Lastly, we find that above primary school education helps to completely offset potential negative effects of flooding.

1 Introduction

In the future, global climate change is expected to cause severe adverse effects on human development (IPCC 2014). In particular, rising temperatures and an intensifying hydrological cycle will likely lead to further increases in the frequency and severity of extreme weather events and natural disasters around the globe (Rummukainen 2012). One of the countries most affected by climate change is Thailand, currently ranking 9th in terms of the long-term climate risk index provided by Germanwatch (2018). Due to its topography, Thailand is already today prone to natural disasters, especially droughts and flooding. In 2011, the country attracted global attention when experiencing one of the worst floods to date (Noy et al. 2019). The disaster affected 13 million people, caused 813 deaths and lead to reparation and rehabilitation costs of US$46 billion (Guha–Sapir et al. 2012).

As the frequency of such disasters intensifies in Thailand (Marks 2011), it becomes more critical to understand how vulnerable households are affected by flood events and how they can cope with them. Understanding household level effects is in particular important in order to derive effective risk mitigation strategies and relevant policy implications. Several studies aim to quantify the impact of flooding on households in South East Asia (Banerjee 2010; Gröger and Zylberberg 2016; Chantarat et al. 2016; Noy et al. 2019). Most of them do so in an event study framework, analyzing the effects of a single flood event on income or expenditure in subsequent years. Fewer studies focus on the effects of recurrent disasters and potential household adaptation to such shocks (Lertamphainont and Sparrow 2016; Felkner et al. 2009; Anttila-Hughes and Hsiang 2013).

This thesis extends the existing evidence on household level impacts of flooding in Thailand. We do so by leveraging the Thailand Vietnam Socio-Economic Panel data (TVSEP) which was collected in the time period from 2007 to 2016 and provides detailed information on socio-economic conditions of households located in the rural areas of northeastern Thailand. Complementing previous research, we combine this unique panel data set with objective flood measures derived from satellite data to estimate impacts of recurrent floods in a multiyear setting. Exploiting random spatio-temporal fluctuations in flood intensities we estimate causal impacts of flooding on household income and expenditures. Furthermore we explore and compare shock coping strategies and identify household level differences in flood resilience.

Doing so, we address multiple strands of the literature. Firstly, we add to the literature employing objective, satellite based measures to quantify the economic impacts of natural disaster (Elliott et al. 2015; Bertinelli and Strobl 2013; Mohan and Strobl 2017) and in particular flooding (Guiteras et al. 2015; Gröger and Zylberberg 2016; Kocornik-Mina et al. 2016; Chen et al. 2017). Subjective, self-reported data was often found to be inappropriate to reflect the true severity of flooding, underlining the need for an objective measure (Guiteras et al. 2015).[1]
Secondly, by proposing two different objective flood measures which distinguish between the absolute flood severity and the relative flood severity on a village level, we contribute to the literature on adaptation to flood events (Banerjee 2010; Chantarat et al. 2016; Islam et al. 2018). Employing these two measures we can identify which type of shocks are more relevant in explaining household level effects: flooding which is extreme in an absolute sense or flooding which is extreme in a relative, village specific manner.
Thirdly, we shed light on the causal impacts of flooding on the micro-economic level (McCarthy et al. 2018; Arouri et al. 2015; Gröger and Zylberberg 2016). In particular we add to literature

[1] For a good overview on the usage of "objective" meteorological and climate data in economics, see Donaldson and Storeygard (2016).

that analyzes household level effects of flooding in Thailand (Garbero and Muttarak 2013; Noy et al. 2019).

We find that flooding in the rural areas of northeastern Thailand has severe negative impacts on household income. Our results suggest that adverse effects are rather driven by deviations from village specific, median flood exposure than by absolute levels of flooding. This indicates a certain degree of adaptation to flooding on the village level. Notably, flood events which lie above the village specific median flood exposure, reduce per capita income by about 10%. Farming income is particularly affected and reduces by about 12.5%. For extreme flood events, 1.4 standard deviations above the village specific median, farming income drops on average by 23%. Simultaneously household expenditures for food and health products rise significantly.
First analyses of potential mitigation channels identify the following mechanisms for income smoothing: The share of households which receive remittances or public transfers increases after flood events. However, these safety nets appear to be insufficient to compensate for the severe losses that households experience. With respect to resilience building, we show that households with better educated household heads are able to completely offset potential adverse effects of flooding. However, the majority of household heads in our sample (89%) only received primary education. This emphasizes the importance of increasing the over all education level in rural Thailand.

The remainder of the paper is organized as follows. In section 2, we review and summarize previous literature on related topics, emphasizing not only household level impacts of natural disasters but also the methodology used to generate accurate and objective impact measures. In section 3, we introduce the data employed in our analyses. We describe the TVSEP panel data set and explain the construction of the two objective flood measures derived from satellite data. The methodology to derive causal estimates of flood impacts is explained in section 4. In section 5 we present our results and conclude with a discussion in section 6.

2 Literature Review

There is currently an extensive line of research that examines the negative impact of flood events on various economic outcomes of households in South East Asia (see for example Poapongsakorn and Meethom 2012, Arouri et al. 2015, Noy et al. 2019). In order to estimate the true causal impact, many studies take an event study approach focusing on the effect of a single flood event. A study by Noy et al. (2019), for example, analyzes the effects of the 2011 flood using a differences in differences analysis. The authors combine panel self-reported data from the Thai Household Socio-Economic Survey (THSES) with satellite data and find a negative effect of the flood on economic outcomes of directly and indirectly affected households. In another event study, Chantarat et al. (2016) analyze the effects of the discontinuity generated by the same 2011 flood in Thailand on changes in preferences, subjective expectations, and behavioral choices among rice-farming households. Based on the results, the authors make a number of policy recommendations including designing safety nets to help agricultural households in the aftermath of a flood, appropriate public insurance policies and development programs that facilitate risk sharing between more flood-prone and less flood-prone households.

To our knowledge there is currently limited literature analyzing the impact of recurring flood events in a panel setting[2]. One similar study to ours conducted by Anttila-Hughes and Hsiang (2013) exploits annual variation in typhoons in the period between 1979-2008 to analyze the

[2]As a reference, please see a paper by Dell et al. (2012) that provides a detailed overview of a rapidly growing body of literature on research using panel data to examine the impact of natural disasters on economic outcomes

average effect of a prior year's typhoon on an average household in the Philippines. The authors find that a local exposure to a typhoon reduces household income by 6.6% and expenditures by 7.1% and leads to substantial increases in infant mortality, comprising 13% of the overall infant mortality rate in the Philippines. Thus, our paper contributes to the current literature by utilizing a panel setting to determine the average causal impact of the precedent year's flood event on households in Thailand over the period of 10 years between 2007 and 2016.

In order to construct the objective flood measure we leverage the methodology used in the paper by Gröger and Zylberberg (2016) that uses the variation in the severity of inundation caused by the 2009 typhoon in Vietnam as an instrument for an exogenous income shock and evaluates the response of affected households with respect to migration. They do so by leveraging satellite data to construct an objective impact measure for the typhoon damages. Employing a large and detailed survey data set combined with an objective impact measure, they analyze household behaviour in the aftermath of a disaster as well as the effectiveness of internal migration to cope with losses. In this sense, they use the household survey data from the Thailand Vietnam Socio Economic Panel (TVSEP) that we also use in our study for Thailand.

In contrast to many other studies Gröger and Zylberberg (2016) do not rely on a self-reported impact measure but construct an objective measure from satellite data to quantify the destruction caused by the typhoon. The motivation for using satellite data in this context is twofold. First, it allows to compute impact measures if on-the-ground data is lacking. Second, it should reduce the issue of systematic biases in the data. Indeed, Guiteras et al. (2015) observe that households do not recall flood events very accurately, especially when floods are frequently experienced by the household. This highlights the importance for an objective measure.

Inundation Measure

An extensive body of research primarily relies on rainfall data in order to approximate local inundation, see for example Yang and Choi (2007) and Roggemann (2015). Paxson (1992) conducts a study on the response of savings to transitory income in Thailand using time-series information on regional rainfall. In the study, households are matched to the weather station that was closest to their district. However, given that there were only 61 weather stations in 1992 in Thailand, this measure does not serve as a good proxy to provide precise results. At the same time, rainfall is only an indirect measure since it is only one of the three sources of flooding, the other two being cyclones and river overflow (Brammer et al. 1990). For this reason, rainfall does not necessarily affect solely the region it was measured in, but might be dragged along by rivers and cause destruction in other regions (Chen et al. 2017; Noy et al. 2019). There are also some studies that consider the upstream water balance to measure inundation, such as the study by Chaney (2013) that uses the height of the Nile to measure flooding in Egypt. However, for the purposes of our study, such a measure would be incomplete, since we need to capture overall inundation including rainfall. The more recent studies rely on geospatial data collected by satellites (Xu 2006; Sakamoto et al. 2013; Sayama et al. 2017; Chen et al. 2017). Xu (2006) constructs the Modified Normalized Difference Water Index (MNDWI) based on the Normalized Difference Water Index (NDWI) by Gao (1996) to distinguish between water and non-water features based on surface reflectance. The cheap availability of high-resolution satellite data has become a powerful tool in economics, which is what we employ to calculate the flood index in our study.

One drawback of using an objective measure is that it does not necessarily take into account that people frequently adapt to the external conditions, thus a household in a more flood prone area may view a flood of the same magnitude differently compared to a household that is not as frequently exposed to flood events (Guiteras et al. 2015). This adaptation hypothesis is further supported by a study from Felkner et al. (2009) that analyzes the impact of climate change on rice production in Thailand over a five year period under the low and high greenhouse gas emission scenarios. Their findings show that farmers are able to adapt to the mild effects of climate change and even benefit from low levels of rainfall. However, this result is not as evident for poorer households with income below the median level. At the same time, both income groups are unable to offset the adverse effects of more extreme changes in weather. Given the adaptation hypothesis, our paper contributes to the literature by also considering the objective flood measure relative to what the households experience in a "normal" flood year.

Despite the number of flood events in Thailand in our sample period, the region is also extremely vulnerable to droughts in particular with respect to agricultural outputs and crop yields as a large share of population is engaged in farming activity as their main occupation. Thus, in the same year, too little rainfall can be an indicator of drought and hence crop failure and clean water shortages, whereas too much rainfall can cause flood disasters[3]. Indeed, one study by Banerjee (2010) that analyzes the impact of floods in Bangladesh argues that there is actually a positive relationship between monsoon floods and agricultural performance. A certain level of flooding seems to be beneficial as it facilitates rice cultivation in the area. In general, the results show that more flood prone areas have higher yields than less flood prone areas. However, this relationship is reversed in "extreme" flood years. Drought years, by contrast, lead to an acute deficit in productivity, specifically a 5 percent shortfall in rice production.

Given the persistent adverse effects of floods and droughts that Thai households experience each year, a question arises on how such households improve their resilience in the aftermath of a shock. A study by Islam et al. (2018) shows that individuals who have family members working abroad receive remittances to assist them during the times of disaster. Gröger and Zylberberg (2016) find a similar result. In terms of education and gender of the household head, the current evidence is not as clear. One study by Anttila-Hughes and Hsiang (2013) shows that male headed households with maximum primary education suffer slightly larger income losses, however not significant at an accepted level. Thus, in our study we further explore such household characteristics to measure their effect on the resilience to the shock.

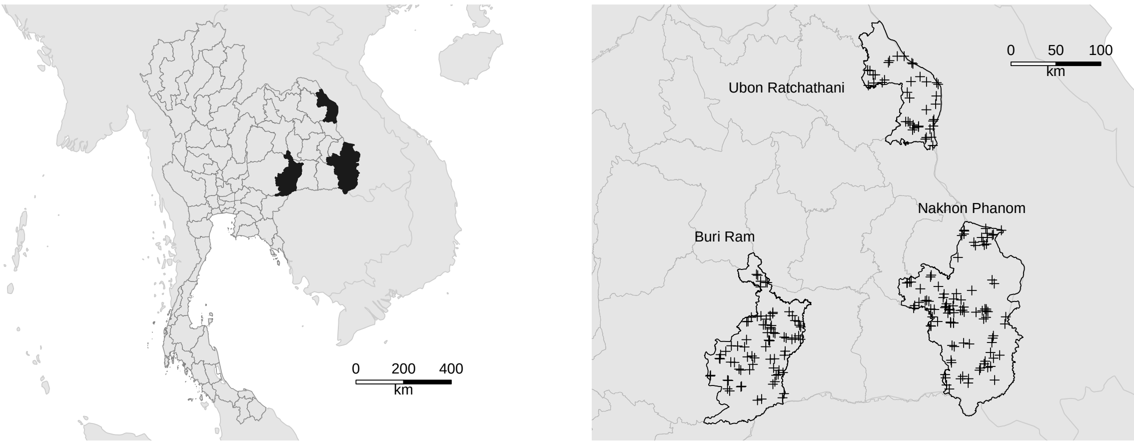
Based on the literature laid out above we derive the following main hypothesis: First, households that experience a severe flood shock will see a decrease in their income while households which experience average levels of flooding will not or only be mildly affected by a flood shock. Secondly, households that receive remittances from friends or relatives will be able to smooth their consumption due to these remittances. Lastly, given inconclusive evidence on household resilience to shocks, we further hypothesise that households with higher levels of education experience less negative impacts from flooding.

Our research objectives are hence twofold:

1. Identify the causal impacts of flooding on income and expenditures of rural households in Thailand.

2. Explore and compare shock coping strategies and identify household level differences in flood resilience

[3] `https://asiancorrespondent.com/2016/10/thailand-breaking-cycle-flooding-drought/`



(a) Location of survey provinces in Thailand **(b)** Locations of villages within survey provinces

Figure 1: Overview of survey provinces and village locations

3 Data

3.1 Household Data

Our analysis is based on household micro-level data from the Thailand Vietnam Socio-Economic Panel (TVSEP) funded by the German Research Foundation (DFG). The study uses six panel data waves conducted in the rural areas of three provinces in Northeastern Thailand (Buriram, Ubon Ratchathani and Nakhon Phanom) in 2007, 2010, 2011, 2013 and 2016.

The TVSEP survey collects socio-economic indicators. These relate back to the previous year asking for example about the household's income. All three of the study provinces feature agricultural employment shares of more than 60%, this is exemplified by the share of agriculture in total GDP, which accounts from 16% in Ubon Ratchathani to 21% in Buriram and 22% in Nakhon Phanom, in contrast with the national average of 9%. A map of the survey provinces is presented in Figure 1.

Within the three provinces surveyed, 2.186 households have been randomly sampled to participate in the survey, 980 in Ubon Ratchathani, 820 in Buriram and 409 in Nakhon Phanom. Over the years, the number of survey households reduced slightly to 1.936 households in 2016. This corresponds to an attrition rate of 11.44%. Moreover, it has to be noted that in 2011 the survey was only carried out in Ubon Ratchathani which effectively reduced the number of observations to 916 in that year. The reason for only sampling one province in 2011 is somewhat unclear, but it seems as if this survey round was conducted more or less spontaneously i.e. out of the original schedule.

Table 1 shows some further details of the composition of the sample used in our analysis. Note here, that all values were computed using a cleaned dataset in which some observations were dropped due to obvious false reporting, missing values or implausible values. Further information on the data cleaning process of the household data can be found in Appendix B.

The sample used for our analysis comprises 9,691 observations over all waves in all regions. Among the three surveyed regions, the largest share of observations come from Ubon Ratchathani, the largest province of the three, followed by Buriram and Nakhon Phanom. Taking a closer look at the household demographics reveals that the three sample regions are quite homogeneous with

Table 1: Summary statistics

Sample	All	Buriram	Ubon Ratchathani	Nakhon Phanom
Observations	9,691	3,411	4,613	1,667
Household Demographics				
Number of men (16-59)	1.76	1.71	1.79	1.78
	[1.73 1.78]	[1.67 1.74]	[1.75 1.82]	[1.72 1.83]
Number of women (16-59)	1.71	1.64	1.71	1.73
	[1.67 1.714]	[1.61 1.68]	[1.68 1.75]	[1.68 1.79]
Number of children (0-14)	1.04	1.03	1.04	1.07
	[1.020 1.062]	[0.99 1.06]	[1.01 1.07]	[1.01 1.12]
Household Head				
Main occupation: farmer	0.61	0.61	0.59	0.66
	[0.60 0.62]	[0.59 0.63]	[0.58 0.60]	[0.64 0.69]
Age	57.60	58.34	57.39	56.69
	[57.35 57.86]	[57.90 58.77]	[57.02 57.77]	[56.08 57.30]
Maximum primary education	0.89	0.88	0.89	0.91
	[0.881 0.894]	[0.87 0.89]	[0.88 0.89]	[0.90 0.93]
Female household head	0.29	0.28	0.29	0.33
	[0.28 0.30]	[0.27 0.30]	[0.27 0.29]	[0.31 0.35]
Household Income				
Total income	7,339	7,182	7,871	6,195
	[7,168 7,511]	[6,907 7,458]	[7,602 8,141]	[5,860 6,530]
Farming income	1,730	1,485	2,089	1,238
	[1,648 1,812]	[1,341 1,629]	[1,968 2,210]	[1,073 1,402]
Self-employed income	1,216	1,008	1,449	998
	[1,109 1,324]	[854 1,162]	[1,271 1,627]	[775 1,221]
Labor income	2,063	2,065	2,218	1,634
	[1,979 2,149]	[1,934 2,195]	[2,078 2,359]	[1,486 1,782]
Formal public transfers	215	204	248	149
	[202 229]	[181 226]	[227 270]	[135 163]
Transfers from relatives	793	935	644	917
	[753 833]	[859 1,010]	[595 692]	[807 1,027]
Household Expenditure				
Total expenditure	6,536	6,079	6,987	5,656
	[6,330 6,743]	[5,726 6,431]	[6,696 7,278]	[5,183 6,129]
Food expenditure	1,675	1,689	1,730	1,506
	[1,650 1,700]	[1,650 1,727]	[1,692 1,770]	[1,448 1,565]
Nonfood expenditure	759	782	774	671
	[748 771]	[764 801]	[757 791]	[647 695]
Health expenditure	99	94	99	113
	[96 104]	[88 101]	[94 104]	[102 123]
Education expenditure	256	294	239	227
	[244 267]	[273 314]	[222 254]	[202 251]
Household finance				
Total savings	1,062	914	1,144	1,096
	[1,018 1,106]	[841 986]	[1,082 1,205]	[973 1,219]
Total borrowings	3,090	3,265	3,035	2,869
	[2,991 3,188]	[3,093 3,437]	[2,898 3,172]	[2,628 3,110]

Notes: 95% confidence intervals in parentheses. All monetary values are given in 2005 PPP USD.
Dependents are defined as children of 14 years old or younger since people older or equal to 60 might still be in employment resulting in a noisier measure.

regard to the number of male and female household members of working age, i.e. 15-59 years old, as well as the number of children, i.e. people younger than 15 years old. This is consistent with the average composition of the households in the northeastern region and Thailand as a whole.[4] In this regard, there is approximately a 50-50 split between number of men and women in the population. The share of children aged between 0 and 14 years old is approximately 34%.

A striking detail is the large share of household heads who stated that farming is their main occupation. This share is highest in Nakhon Phanom with nearly 66% and lowest in Ubon Ratchathani with around 59% . Indeed, according to the nation wide statistics between 2014 and 2017, almost half of all farm households of Thailand were located in the Northeast region with Ubon Ratchathani having the majority (approximately 9%) compared to the three provinces under consideration. Also, the share of household heads whose highest completed education is primary school is high with an average of around 89%. The TVSEP data set defines primary education as the first 7 years of schooling. Again, this share is highest in Nakhon Phanom and lowest in Ubon Ratchathani. Looking at the nation-wide statistics, people of working age, i.e. 15-59 years old, have on average 9 years of schooling, which is more or less consistent with the household data under consideration.

Before discussing the yearly household income in our sample it has to be noted that Thailand is characterized by strong spatial heterogeneity with respect to income. While the Bangkok area as well as the South are relatively rich, the Northeastern region of Thailand is much poorer [5]. While the per capita GDP between 2009 and 2015 in all Thailand was around 16,178 USD (2005 PPP adjusted), the mean household income in our sample was only around 7,339 USD. Divided by an average household size of 4.5 household members, this only amounts to a yearly per capita household income of 1,631 USD.

Although the majority of household heads in our sample regions seem to work in agriculture, this does not directly translate into the sources of household income. While around 60% of the household heads work in agriculture, only around 18% of household income is generated from agriculture, while around 47.5% come from non agricultural labor income. One potential explanation could be that many of the farmers live at subsistence level, therefore even though they work as farmers, they do not get any income. To check this hypothesis, we have computed the share of agricultural production which is consumed by every household. The results are clear, 45% of the household consume all they produce, moreover 60% of the households consume more than 50% of their production.

With regard to income, the survey regions are somewhat more heterogeneous with Ubon Ratchathani showing the highest level of household income of 7,871.176 USD (2005 PPP adjusted) and Nakhon Phanom households having the lowest income of 6,195.022 USD. Based on the data collected by the National Statistical Office (NSO), average annual income per household between 2007 and 2017 is higher than estimated in the TVSEP data set. Ubon Ratchathani still shows the highest level with 15,081 USD (2005 PPP adjusted), followed by Buriram with 11,400 USD and Nakhon Phanom with 10,404 USD. This stark difference is likely because the TVSEP survey only collected data for rural households while the NSO data also considers urban areas. Lastly, remittances from relatives seem to play a significant role in household income taking up a share of slightly less than 10%. Again, there is quite some heterogeneity across regions with this share being around 15% in Nakhon Phanom and only 8% in Ubon Ratchathani.

[4]Data provided by National Statistics Office of Thailand (NSO) and retrieved from `http://statbbi.nso.go.th/staticreport/page/sector/en/index.aspx`, last accessed 2/6/2019

[5]For further details see `http://www.thaiwebsites.com/provinces-GDP.asp`, last accessed 2/6/2019

A look on the expenditure side of household finances reveals that households spent on average 6,536 USD (2005 PPP adjusted) per year. Out of all expenditures, food expenditures account for roughly 25%, a share that is more or less stable across all survey provinces. Similar to the estimates of household income, the expenditure estimates are significantly lower compared to the data obtained from the NSO, which reports the average household expenditure at around 10,000 USD. Lastly, table 1 shows that households in our sample saved on average 1,062 USD. The savings rate is highest in Ubon Ratchathani, the richest province, where households save on average around 15% of their income or 1,143 USD in absolute terms. Buriram, on the other hand has the lowest savings rate of around 13%, or 913 USD. Thus, the households with the highest income level out of all three provinces have the highest savings rate. Total borrowings of households, on the other hand, are highest in Buriram with on average 3,264 USD and lowest in Nakhon Phanom with 2,868 USD. About 50% of the borrowed resources are spent on household consumption across the three provinces.

Given that an important building block of this study is the actual size/significance of the shock, i.e the recurrent floods, we are also interested in how the survey households perceive such shocks and how this compares with other relevant shocks faced by them. In this respect, flooding has been reported as the most dramatic shock faced by around 10% of households on average. Together with droughts, these two natural shocks are reported by a range from 22% to 40% of the households.

3.2 Objective flood measure

A central component in our analysis is the generation of an objective measure which captures the severeness of flooding on the village level. For this purpose we combine detailed GPS data on village locations with annual and monthly flood maps derived from satellite imagery. We rely on two different sources for satellite based flood data:

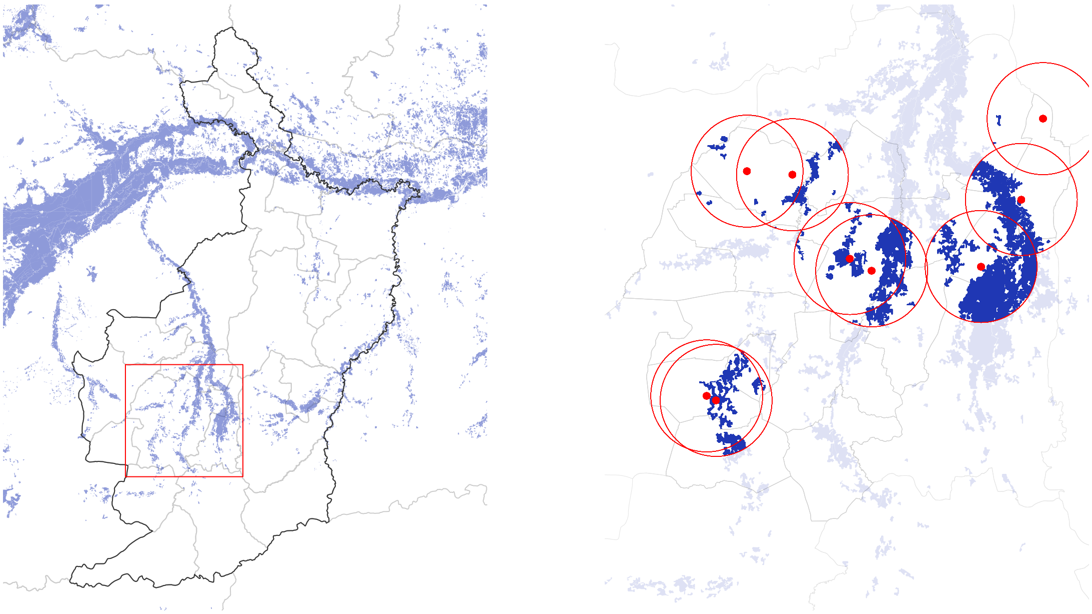
1. Annual flood maps obtained from the Thailand Flood Monitoring System, a project of the Geo-Informatics and Space Technology Development Agency (GISTDA) of Thailand
2. Monthly flood maps obtained from the United States National Aeronautics and Space Administration (NASA)

In this thesis, we focus on the annual flood maps provided by GISTDA. They are available on a finer spatial but lower temporal resolution than the NASA data.

GISTDA flood maps

For the generation of detailed maps of flooding in Thailand, GISTDA evaluates satellite images from various satellite systems including Radarsat, Cosmo SkyMed and Thaichote. All systems produce daily imagery of parts of Thailand at different, but fine spatial resolutions ranging from 2m to 100m. GISTDA uses various surface water detection algorithms to extract information about the extend of flooding in areas where satellite imagery is available. The resulting high resolution flood maps, which only cover parts of Thailand, are merged over multiple days to obtain a complete overview of a flood event. A selection of the resulting maps containing the extent of single flood events is available on the GISTDA flood monitoring website.[6]

[6] When this thesis was written, two versions of the website were active:
`http://flood.gistda.or.th/indexEN.html`, last accessed: 23/05/2019
`https://floodv2.gistda.or.th/`, last accessed: 24/05/2019.



(a) 2010 flooding in Buri Ram and surrounding provinces. Red lines mark the location of the Nong Rong district.

(b) Detailed overview of flood index construction

Figure 2: Overview of flood index construction for villages in Nang Rong district of Buri Ram. Panel (a) displays the 2010 flood for the region as well as the area in which the villages lie. Panel (b) shows the exact location of each village, red dot, as well as the 5 km area around each village, red circle. The flood index $FL_{v,2010}$ is calculated as the percentage flooded area within the red circles, here displayed in dark blue.

We use aggregated information for the construction of our annual flood measures. GISTDA provides annual flood maps which result from merging all available flood maps of single flood events that occurred in a particular year. Subsequently, the annual data provides binary information on whether a flood occurred at a given location at least once within a year. Any information on the duration of a flood or on whether a location was flooded multiple times within a year is lost in the merging process. This is the key weakness of the GISTDA annual data. To get a visual impression of the data, an exemplary, complete flood map of Thailand for the year 2010 is provided in appendix A.

An objective measure for absolute flooding

For our analysis we use annual flood maps for the years 2006 to 2016.[7] Based on these maps, we calculate a simple objective annual flood measure, denoted FL, for every village v in the TVSEP data. We define the measure $FL_{v,t}$ as the percentage area flooded within a 5 km radius around village v at least once in year t:

$$FL_{v,t} = \begin{array}{l}\text{percentage area flooded within a 5 km radius around village } v \\ \text{at least once in year } t\end{array} \quad (1)$$

Figure 2 illustrates how the index is constructed for a set of exemplary villages that lie in the Nang Rong district of Buri Ram. Panel (a) displays the 2010 flood data for Buri Ram and surrounding provinces. Panel (b) zooms into the area of the Nang Rong district and shows the

[7] A map for 2005 was also available but showed clear irregularities and large amounts of missing data. It was therefore excluded.

location of the respective villages (red dots), as well as, the 5km buffer zone around each village (red circle). The flood index $FL_{v,2010}$ is calculated as the percentage area flooded within each circle (highlighted in dark blue).

We choose a radius of 5km for the construction of the flood measure since TVSEP data from 2007 indicates that 90% of all farming activities occur in a 5km radius of household locations. Therefore our flood measure is specialised on measuring effects of flooding on local farming activities. Nevertheless it might also capture other disruptions, for example if infrastructure close to the village is damaged by floods. Gröger and Zylberberg (2016) construct their inundation measure in a similar manner and also use a 5 km radius. For robustness checks we used alternative radii of 10km, 20km and 50km.[8]

An objective measure for relative flooding

Motivated by previous research such as Guiteras et al. (2015) or Sirisankanan (2016), we hypothesize that people adapt to the "normal" flood exposure conditions. To measure effects of flooding beyond the village specific normal conditions, we propose a second, alternative flood index, denoted $FD_{v,t}$. Differently to the first index, this measure is a standardized, village dependent index which captures deviations of flood exposure from normal years. It is calculated as the scaled deviation of median flood exposure:

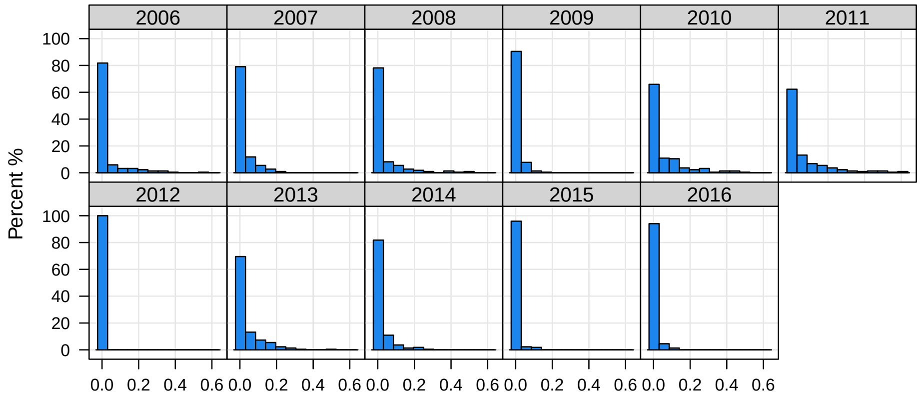
$$FD_{v,t} = \frac{FL_{v,t} - \text{median}\left(FL_{v,2006:2016}\right)}{\text{sd}\left(FL_{v,2006:2016}\right)} \tag{2}$$

Here, we use the deviation from the median flood intensity as opposed to the average flood intensity to be more robust to outliers. Furthermore we divide by the standard deviation to obtain a measure which depends on the village specific yearly variation in flooding. If the index takes for example the value $FD_{v,t} = 2$ this implies that in year t village v experienced a flood which was 2 standard deviations above the village specific median flood exposure. The procedure implies that a village, which usually experiences flood exposure close to its median, obtains a very high FD index in a year of an extreme flood. By contrast, a village which usually experiences large fluctuations in flood exposure across years will show a comparably lower FD index even in a year of an extreme flood.

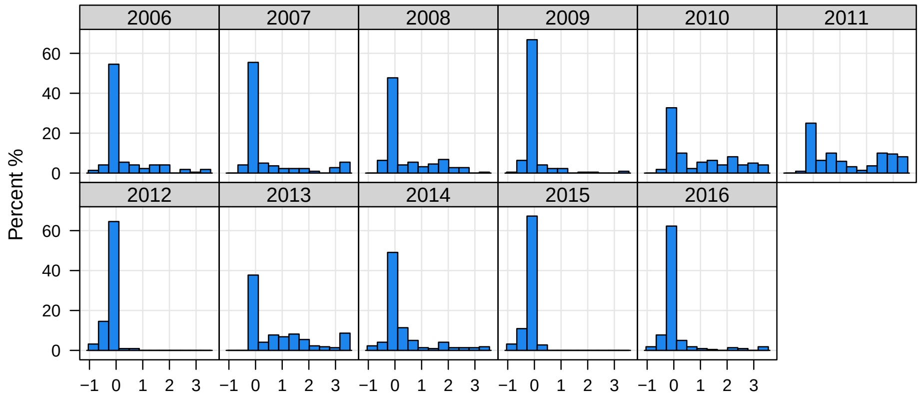
Figure 3 shows the distribution of the two different village level flood indices for all years from 2006 to 2016. Panel (a) shows the distribution of the FL indicator and reveals that the majority of villages in the TVSEP panel only experience very little flooding within a 5km radius. With respect to our FL measure, every year about 80% of the villages experience basically no flooding. Only in the years 2010, 2011 and 2013 more than 30 % of the sample villages were affected by flooding. Especially in 2011, the worst flood year for Thailand on record, some villages experienced floods which inundated more than 50% of their surroundings at least once. By contrast, in 2012 not a single village was hit by a flood and also in 2015 less than 10% of the villages in our sample were effected by flooding. Both years were particularly severe drought years.

Panel (b) in figure 3 shows the distribution of the standardized flood deviation index FD. The index allows us to analyse the flood severity in relative terms. It still takes the value 0 frequently but shows more variation than the FL index and emphasises the severeness of the 2010, 2011 and 2013 floods: For example in 2011 about 30% of the villages in our sample experienced flooding more than 2 standard deviations above their median flood exposure. In a similar way in which the FD measure emphasises severe flooding, it might also be able to capture effects of relative

[8] Figure A.2 in the appendix displays the distribution of the FL indicators for different years



(a) Simple flood intensity index FL



(b) Flood deviation index FD

Figure 3: Overview of the distribution of the villages level flood indices by year. Panel (a) shows the distribution of the FL index defined as the percentage area flooded at least once within a 5km radius of a village. Panel (b) displays the distribution of the village dependent flood deviation measure FD defined in equation 2.

local droughts. For example in 2012, a drought year, about 30% of the villages have a below median flood exposure. The relationship of flooding and droughts is of central concern for our analysis, as we aim to causally identify the effects of flooding. We will explain further details on our identification strategy in the methodology section below.

Robustness check - subjective flood measure

In order to assess whether the objective inundation index is a sensible measure for gauging the impact of flooding on households, we derived a subjective flood measure from data included in the TVSEP. Running a correlation analysis between our objective measure and the subjective measure yields mixed results. This is in line with Guiteras et al. (2015) who show that subjective measures are often a weak proxy for true flood exposure. A more detailed discussion regarding the way we generate the subjective measure and the correlation between the two measures can be found in the Appendix C.

3.3 Additional Data

Besides the TVSEP panel data and the flood maps provided by GISTDA and NASA, we utilize monthly rainfall data as a control in our regressions. The data was obtained from the Tropical Rainfall Measurement Mission (TRMM), a project coordinated by NASA. In particular we leverage the TRMM 3B43 Version 7 data product, which merges satellite rainfall estimates with ground gauge measurements into gridded "best" precipitation estimates on a calendar month temporal resolution and a 0.25° by 0.25° spatial resolution. For Thailand this roughly equals a resolution of about 26km × 26km. Data was obtained for the area of Thailand for all months between January 2005 and December 2016.

To calculate a village specific, monthly rainfall index RF from the TRMM data, we proceed in a similar manner as for the flood index. For each village v and each month m in year t, we calculate the average precipitation in an 5km radius around the village location:

$$RF_{v,t,m} = \begin{array}{l}\text{average precipitation in a 5km radius around village } v \\ \text{in month } m \text{ of year } t\end{array} \tag{3}$$

Since the gridded rainfall data has a rather coarse resolution, this implies that for most villages we simply assign the average monthly rainfall estimate of the grid cell the villages is located in.

4 Methodology

Our study has two main objectives:

1. Identify the causal impacts of flooding on income and expenditures of rural households in Thailand.
2. Explore and compare shock coping strategies and identify household level differences in flood resilience

In the following we describe the methods and identification strategies used to address both objectives.

Temporal alignment of household and flood data

Our first concern is the misalignment of household data and flood maps: Each of the 6 TVSEP panel rounds was conducted between April and June of the respective survey year. Our flood indicators FL and FD are however only available on an annual basis and measure flooding for the entire calendar year starting in January and ending in December. However, we know from several reports (see for example Noy et al. 2019; Lertamphainont and Sparrow 2016), as well as, from the monthly NASA data that flooding in Northeastern Thailand mostly occurs during the months of August to November, following the monsoon season which starts in June and ends in September. At the beginning of the year, our survey provinces only experience little rainfall and flooding is not an issue. Furthermore, most of our survey households heavily rely on farming of wet rice, i.e. jasmine rice, and harvesting calendars for Thailand indicate that growing periods for this rice coincide with typical flooding periods in autumn (United States Department of Agriculture). If the rice harvest is affected by flooding, we expect negative impacts on the household level. For this reason we analyse the impact of *previous year flooding* on household income and expenditure. For example this means that we are interested in the effect of our FD indicator for the year 2006 on the annual income of a household measured in spring 2007. Figure 4 shows the timeline of the individual survey rounds and the flood seasons.

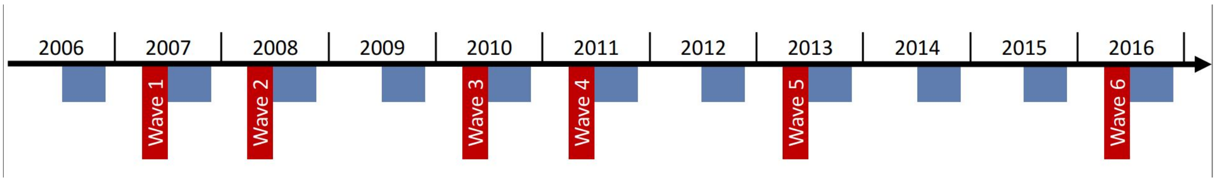


Figure 4: Timeline of the TVSEP survey round. Red boxes indicate the survey period, blue boxes indicate the monsoon season. Flood periods usually proceed the monsoon season.

Non-random exposure to flooding

A major threat to identification is the non-random exposure of households to flooding. If exposure to flooding on the village level is correlated with other factors that explain household income or expenditure, causal identification of the effects of flooding is not feasible due to omitted variable bias. Obvious time invariant factors, that could be correlated with flood exposure but simultaneously explain household income and expenditures, are for example soil fertility in the surrounding of a village, a village's proximity to rivers or its other geographical characteristics.

Time variant factors that could affect both flooding as well as income and expenditures are mainly meteorological in nature. In particular we are concerned with droughts and rainfall, which are correlated with flooding. Here, it is important to point out that the correlation structure might not be as obvious as it seems on first sight. While in a short time period, e.g. a month or a week, droughts and floods are very unlikely to coincide, this might not be the case on an annual time scale. In 2010 Thailand experienced for example extreme droughts previous to the monsoon period as well as extreme floods towards the end of the year (Garbero and Muttarak 2013). Hence flooding might or might not coincide with droughts. With respect to rainfall, Guiteras et al. (2015) and Chen et al. (2017) show that rainfall is positively correlated with flooding but may be an imperfect proxy for floods, given factors such as upstream water balance, proximity to rivers, and topography.

Identification Strategy

To identify causal impacts of flooding it is important to control for both, time variant and time invariant confounding factors. In our main regression specification we therefore use a panel fixed effects model of the following form:

$$y_{pvht} = \mu_h + F_{vt-1}\beta + X_{vht}^T\gamma + \delta_{pt} + \varepsilon_{pvht} \tag{4}$$

Here p indicates the province, v the village, h the household and t the year. y_{pvht} is the respective dependent variable of interest; in our case mainly income and expenditure variables. F_{vt-1} is the village level flood measure, either the absolute version FL_{vt-1} or the village dependent relative measure FD_{vt-1}, and β the main coefficient of interest. Note that we analyse the effect of the previous year flood, $t-1$, on annual income and expenditure measured in year t.

μ_h is a household fixed effect, which captures time invariant household characteristics including location specific characteristics. Hence the household fixed effect is crucial for our identification strategy as it allows us to capture time invariant confounders, potentially correlated with flooding, income and expenditures.

To control for time variant confounders we include a province year fixed effect δ_{pt}, as well as, a vector of additional controls X_{vht}. By including a province year fixed effect we intend to capture part of the correlation between droughts and floods. Droughts are a more regional phenomena than floods and the province year fixed effect is supposed to capture this regional variation. To ensure that our results are not confounded by any additional effects of droughts or floods, we include a vector of previous months rainfall averages in our controls. In particular X_{vht} contains linear and squared monthly rainfall data (RF) for all 12 months preceding the start of the respective survey in April. For example, for data from the survey round in $t = 2010$, $X_{vh,2010}$ includes all linear and squared rainfall averages for village v, for the months April 2009 to March 2010.

Besides the necessary controls which are crucial to address confounding factors, we also include controls for household size, age and sex of the household head in all of our regression. These factors should not confound our regressions but might change over time and potentially absorb additional variance in the respective variable of interest y.

A crucial aspect for all of our regressions is the correlation of standard errors ε_{pvht}. Clearly observations for multiple households located in the same village are correlated and therefore assuming independent standard errors would be fatal. However, even allowing for village or district clusters would be insufficient to capture the potentially complex spatio-temporal error structure. For example district clusters would not allow for any correlation between two households from different districts, even if the two households lie spatially very close to each other and are only separated by the district border. For this reason we employ a combination of Newey-West and Conley standard errors in all of our regressions (Newey and West 1987; Conley 1999). For both types of errors we use a Bartlett kernel and allow for temporal correlation up to a lag of 5 years, as well as, for spatial correlation up to 100 km.[9]

Non-linear flood effects

Based on previous literature such as Chen et al. (2017) and Banerjee (2010), we expect that flooding might have non-monotonic impacts on household income and expenditure. Especially extreme events at the top end of the distribution of floods are likely to have very different impacts on households than average flood events which villages are exposed to more frequently. To analyze potential non-linearities and obtain more robust results on the impacts of flooding, we bin each of our 2 flood indices into 4 bins. For the absolute flood measure FL which informs about the percentage area flooded within the 5km surrounding of a village, we chose the following bins:

$$\text{Bins for } FL = \begin{cases} \text{no flooding,} & \text{if } FL = 0 \\ \text{low flooding,} & \text{if } FL > 0, \text{ but does not belong to the top 90\%} \\ & \quad\text{of floods observed across villages} \\ \text{high flooding,} & \text{if } FL \text{ belongs to the top 90\% floods} \\ \text{extreme flooding,} & \text{if } FL \text{ belongs to the top 99\% floods} \end{cases} \tag{5}$$

Note that the 90% quantile of floods with respect to the FL measure is 5%. Hence a flood is classified as a high flood if more than 5% of the area surrounding a village is flooded at least once. The 99% quantile of the FL indicator lies at about 27% of area flooded.

[9] For details on this type of multi-way clustering see Colin Cameron and Miller (2015) and a very good blog entry by Christensen and Fetzer.

Similar to the FL indicator we also bin the data for our village dependent, relative flood measure:

$$\text{Bins for } FD = \begin{cases} \text{below median}, & \text{if } FD < 0 \\ \text{median flood}, & \text{if } FD = 0 \\ \text{above median}, & \text{if } FD > 0, \text{ but does not belong to the top 90\%} \\ & \quad \text{of floods observed across villages} \\ \text{extreme flooding}, & \text{if } FD \text{ belongs to the top 90\% floods} \end{cases} \tag{6}$$

Note that the 90% quantile of the FD indicator is 1.4. Hence all floods are classified as high flood events if they lie 1.4 standard deviations above the median flood level.

Identification of resilient households and shock coping strategies

To address our second research question we proceed in two different manners. To explore whether households compensate potential income losses from flood shocks with remittances and to explore whether they receive additional public transfers after a flood event, we rely on our original regression from equation 4 but use information on remittances and public transfers as the dependent variable y.
For the identification of households that are resilient to flooding we interact the β coefficient in our baseline regression model with information about other household characteristics Z_{pvht}. The resulting regression set up is

$$y_{pvht} = \mu_h + F_{vt-1}\beta + F_{vt-1} \times Z_{pvht}\alpha + X_{vht}^T\gamma + \delta_{pt} + \varepsilon_{pvht} \tag{7}$$

The resulting regression coefficient α allows us to infer whether specific groups of households, as defined by their characteristic Z, react differently to flooding. For example we explore whether households with better educated household heads or households with female household heads are more or less resilient to floods.

5 Results

Before taking a closer look at the detailed results, we determine which of our 4 different indices (FL, FD, FL categorized and FD categorize) is the most appropriate to capture household level effects of flooding. Table 2 gives an exemplary overview for regressing log income per household member on the the four different indices available. Column (1) states the results of using the yearly absolute flood measure, FL, that does not account for village specific median flood exposure. Using it in our baseline regression specification, it turns out to have an insignificant effect on household income. Consequently, we run the regression again using a flood measure based on the deviation from common levels of flooding FD. The negative sign of the coefficient displayed in column (2) suggests that this index is somewhat more sensible, but it is still statistically indifferent from 0. Next, splitting the flood measure into 3 bins allows us to capture potential heterogeneity and non-linearities in the effects with respect to the severity of the flood. Column (3) shows the results of this specification. Now, two out of three coefficients have a negative sign but the one for extreme flood events is positive. Still, all coefficients are insignificant and should not be interpreted. Lastly, we consider the binned measure for village specific deviations of median flood exposure. Using this measure the signs of all three estimates turn negative. In addition the statistical significance of the coefficient for above median levels of flooding becomes highly significant supporting our methodological approach of looking at binned village level deviations from the median flood exposure.

As our inundation measure FD captures the standardized deviations from the median village level flood exposure, all households in villages that showed no deviation from the median level, i.e. households that never experienced a flood, were removed from the analysis. Thus, in table 2 we observe that the number of observations has dropped from 9,691 in columns 1 and 3 to 7,795 in column 2 and 4.

As we find highly significant results when using our binned FD measure we preliminary conclude that in our sample, deviations from median flood exposure rather than absolute flood intensities translate into negative income shocks for rural households. This implies a certain degree of village level adaptation. Therefore, we focus in the remaining section on the results generated when employing the FD measure.

5.1 Household Income

Table 3 presents the results of the main specification 4 for three dependent variables, all measured in natural logarithms and in per household member terms: `log_phm _income` refers to total income, `log_phm_netincome` considers the total income excluding remittances and public transfers and `log_phm_farm_income`, which indicates total farming income.

We find negative coefficients on income per household member although for two of the income measures, negative deviation and top 10% positive deviation from the median flood exposure, this effect is not significant. The estimate for high positive deviations from the standardized median flood exposure, however, suggests that income per household member drops by 10.2%. This result is significant at the 1% level.

In order to get a better understanding about the potential heterogeneous effects on income, we test whether the negative impact of flooding on total income is more severe if households had not received remittances and public transfers. The corresponding results are reported in column (2) of the same table. The number of observations further drops to 7,729 driven by 66 households in our sample that would have had an income below 0 making a log transformation for these values impossible. We observe a similar negative trend as in column (1); however, the point estimates are less in magnitude, which is not what we would expect. In this case, given high positive deviation from the standardized median flood exposure, net income per household member falls by 7.4%, which is significant on the 5% significance. This result is inconclusive and can potentially be explained by selection into receiving remittances or transfers of the dropped observations.

Finally, taking a look at farming income in column (3), the number of observations drops further to 5,718 observations since only 60% of the households receive income from cultivating crops and rearing livestock. We find that for a positive deviation from the median flood exposure, farming income drops by 12.6% on a 10% significance level and by 23.1% on a 1% significance level for an extreme deviation from the median. These results are consistent with the literature that farmers are able to mitigate the adverse effects of floods for the levels of inundation within the first two categories, but not for extreme flood events.

Given the general negative response of the household income levels to flood events, we would like to further analyze whether there is heterogeneity in the magnitude of this effect, in other words, whether households with certain characteristics are more able to cope with the adverse effect of the flood. In particular, we analyze whether it makes a difference if a household head has more than primary education, if a household head is a female and if a household receives any remittances or transfers.

The results for education are presented in Table D.1, which can be found in Appendix D. The coefficient of `1-maxprim` can be seen as a sensibility check, since one can observe a positive and

Table 2: Comparison of different flood indices

	Dependent variable:			
	log_phm_income	log_phm_income	log_phm_income	log_phm_income
	(1)	(2)	(3)	(4)
FL	0.115			
	(0.282)			
FD		−0.018		
		(0.016)		
FL low			−0.020	
			(0.027)	
FL high			−0.022	
			(0.050)	
FL extreme			0.030	
			(0.119)	
FD below				−0.062
				(0.039)
FD above				-0.102^{***}
				(0.032)
FD extreme				−0.071
				(0.052)
Rainfall	Yes	Yes	Yes	Yes
Rainfall^2	Yes	Yes	Yes	Yes
Observations	9,419	7,795	9,419	7,795
R^2	0.479	0.480	0.479	0.481
Adjusted R^2	0.326	0.325	0.326	0.325
Residual Std. Error	0.817 (df = 7288)	0.820 (df = 5998)	0.817 (df = 7286)	0.820 (df = 5996)

Note: This tables shows regression results of the three different flooding indices described in the methodology section. The depend variable in each column is log per household member income. Besides controlling for rainfall and squared rainfall, all specifications control for changes in household size, the gender as well as the educational attainment of the household head.
*p<0.1; **p<0.05; ***p<0.01

Table 3: Effects of flooding on various income variables

	Dependent variable:		
	log_phm_income	log_phm_netincome	log_phm_farm_income
	(1)	(2)	(3)
FD below	−0.062	−0.019	−0.102
	(0.039)	(0.041)	(0.063)
FD above	−0.102***	−0.074**	−0.126*
	(0.032)	(0.037)	(0.070)
FD extreme	−0.071	−0.062	−0.231***
	(0.052)	(0.059)	(0.080)
Rainfall	Yes	Yes	Yes
Rainfall^2	Yes	Yes	Yes
Observations	7,795	7,729	5,718
R^2	0.481	0.482	0.608
Adjusted R^2	0.325	0.325	0.443
Residual Std. Error	0.820 (df = 5996)	0.873 (df = 5930)	1.200 (df = 4023)

Note: This table shows the effect of the three different levels of flooding on different income measures. `log_phm_income` measures income per household member. `log_phm_netincome` measure income per household members excluding remittances and public transfers. `log_phm_farm_income` measures income per household member generated from farming activities including income from both, crops and livestock. Besides controlling for rainfall and squared rainfall, all specifications control for changes in household size, the gender as well as the educational attainment of the household head.
Significance levels are given as follows: *p<0.1; **p<0.05; ***p<0.01

significant relationship between education levels of the household head and per household member income. However, this association cannot be interpreted causally. In general, we observe a negative effect on income if a household head has only maximum primary education. In particular, we observe a 9.9% decrease significant on a 1% level for the high category and 8.9% decrease significant on 10% level for the extreme high category. Net of transfers and remittances, income per household member drops by 8.1% on a 5% level for the high category. Farming income drops by 20.6% on a 5% level for the extreme high deviation from the median flood exposure. In order to compute the effect for a household head with more education, we sum up the point estimates for the flood index with the education interaction term corresponding to the same index. Running a joint significance test on each value for the three income categories yields large p-values indicating that the joint effect is statistically indifferent from 0. Thus, we can conclude that households with better educated household heads are more resilient to flood events.

Similarly, we study the results for the gender of the household head which are presented in Table D.3 in the appendix. Given the results for a joint significance test, in general, we observe that total income decreases more for negative deviations from the median flood if the household is led by a female. This result might be explained by the rather traditional social structure of the Thai society.

Finally, the results for remittances are presented in Table D.2. Since whether household receive remittances or public transfers as well as the volume of such remittances is likely to be endogenous to a flood shock, including these variables as explanatory variables would lead to endogeneity issue. Moreover, both variables are characterized by extremely high levels of heterogeneity as well as many missing and zero values. Thus we the most sensible analysis of remittances and

Table 4: Effects of flooding on expenditure variables

	Dependent variable:			
	TotExp	NFoodexp	Foodexp	healthexp
	(1)	(2)	(3)	(4)
FD below	−0.045	−0.038	−0.017	0.017
	(0.033)	(0.025)	(0.030)	(0.077)
FD above	0.024	−0.013	−0.012	−0.024
	(0.040)	(0.020)	(0.034)	(0.080)
FD extreme	0.040	−0.019	0.071**	0.154*
	(0.036)	(0.027)	(0.033)	(0.093)
Rainfall	Yes	Yes	Yes	
Rainfall^2	Yes	Yes	Yes	
Observations	7,996	7,654	7,305	4,996
R^2	0.573	0.605	0.499	0.502
Adjusted R^2	0.449	0.483	0.336	0.236
Residual Std. Error	0.647 (df = 6194)	0.535 (df = 5853)	0.643 (df = 5506)	1.295 (df = 3260)

Note: This table shows the effect of the three different levels of flooding on different expenditure measures. `TotExp` measures log total expenditure per household member. `NFoodexp` measures log non-food expenditure per household members. `Foodexp` measures log food expenditure per household member. `healthexp` measures log health expenditures per household member. Besides controlling for rainfall and squared rainfall, all specifications control for changes in household size, the gender as well as the educational attainment of the household head.
Significance levels are given as follows:*p<0.1; **p<0.05; ***p<0.01

public transfers can be done by looking at the probability of receiving remittances or public transfers after a flood shock. To analyze this relationship, we run a linear probability model on indicators for both variables. Our results suggest that the effects go in the expected direction, i.e. all coefficients have a positive sign. However non of these effects are statisticall different from 0.

5.2 Household expenditures

Given the general negative impact of flood events on income levels of the households, we expect a similar effect on expenditures. We estimate the main specification 4 including the same controls with the logarithm of household expenditures as the dependent variable and present the results in Table 4. In a similar manner to the household income, the number of observations dropped to $7,996$. For total expenditure per household member, we do not observe any significant results with point estimates pointing in a negative direction for a negative deviation from the median flood exposure and positive direction for a positive deviation. In columns 2-4 we estimate the main regression for different expenditure subcategories, namely non-food, food and health. Non-food expenditures include expenditures on electricity, water, gas and personal care supplies. There is no significant effect on non-food expenditures, however the estimates point to a negative direction, which we expect at a time of a flood.
Food expenditures increase by 7.1% significant at a 5% level for an extreme deviation from the median flood exposure. This is what we would have expected since many households carry out their farming activities on a subsistence level. If their crops are destroyed in a flood, they need to compensate for this loss by buying their produce from a market. Health expenditures include expenditures on medicine purchased from the pharmacy and doctor fees and experience a 15.4%

increase at a 10% significance level after an extreme flood event in the previous year. This finding is again in line with our initial hypothesis since extreme flood events seem to impact households more severely than weaker ones, resulting in higher expenditures for doctors and medicine

Similar to the strategy for analyzing the results for household income, we further investigate household expenditures and determine whether households possess certain characteristics that make them more resilient towards flood events. Our analysis in this regard includes education and gender of the household head. The effect of education on household expenditure outcomes are presented in Table D.4. Those for the sex of the household head can be found in Table D.5. Over all, the results seem somewhat arbitrary with regard to the signs of the coefficients. Moreover all interaction terms are jointly insignificant suggesting that that household expenditures neither depend on the educational attainment nor on the sex of the household head. This is most likely driven by the high level of variation in these observations.

6 Conclusion

The traditional literature on household level impact of flooding has predominantly focused on the effect of individual flood events on various dimensions of the households. One shortcoming of such event studies is that one can neither truly identify the long run effect of reoccurring floods on household level outcomes nor analyze what drives resilience to such shocks. In this paper, we first set out to identify the causal impacts of flooding on income and expenditures of rural households in Thailand. Next, we explored and compared shock coping strategies and identified potential differences in flood resilience based on household characteristics. To do so, we leveraged a detailed household panel data set provided by the Thai and Vietnam Socio-Economic Panel. To quantify the severity of flood events, we calculated flood indices based on flood maps collected by the Geo-Informatics and Space Technology Development Agency (GISTDA) measuring the deviation from median levels of flooding in a 5km radius around a respective village.

Our results suggest a negative relationship between floods and per household member income, for both total income and income from farming. Per household member expenditure, however, does not seem to be affected by flood events at all. The only exemptions are food and health expenditures, which increase after flood events that are among the top 10% of the most severe floods. The former is likely to be driven by the fact that many households in northeastern Thailand live at subsistence level, and therefore consume all of their farming production. A lack of production in a given year may lead these households to substitute this loss by buying produce from markets. Rising health expenditures may be explained by injuries caused or diseases obtained during a heavy flood.

Investigating potential risk mitigation strategies revealed that education seems to be an important channel to better cope with flood events. Generally speaking, the average level of education in the studied areas is low with more than 60% of household members only heaving primary education. This provides valuable insights for policy makers emphasizing the importance of increasing the over all education level in rural Thailand. Moreover, our data suggests that only very few households are insured against potential flood disasters. Unfortunately we were not able to investigate this issue in further detail, due to too much noise in our household data. Thus, this risk mitigation channel leaves some room for future research.

Although our findings are generally in line with what we expected in both signs and magnitudes, they are not particularly robust to alternative specifications or the use of larger radii. The lack of statistical significance as well as robustness of some of our results are likely to be caused by the high level of noise in our household level data as well as potential issues with the flood data used in our analysis. The former is due to measurement error common in survey data. The latter might be due to a lack of accuracy of satellite images as well as only having a flood maps that are accumulated on a yearly level.

The high level of noise in the TVSEP data as well as the over all low level of flooding in the studied areas leads to a low noise signal ratio. A potential solution to this issue is provided by the Townsend Thai Project panel data set which is comparable with the TVSEP data set with respect to the data included, but includes provinces that were flooded more frequently and more severely than those used in this analysis potentially improving the noise-signal ratio. Although we tried to obtain this data set for the purpose of this thesis, we did not manage to do so. Consequently, carrying out this analysis with the Townsend Thai Project data set might provide some more robust insights.

To reduce the issues related to the flood map accuracy, we would suggest to use flood data provided by NASA. These data would allow the researcher to construct flood indices based on monthly flood maps allowing for a more flexible and accurate analysis actual levels of inundation. As mentioned above, we obtained such data but were not able to carry out a detailed analysis due to time constraints. Thus such an analysis is left for future research.

References

Anttila-Hughes, Jesse K., and Solomon M Hsiang. 2013. "Destruction, Disinvestment, and Death: Economic and Human Losses". http://ssrn.com/abstract=2220501.

Arouri, Mohamed, et al. 2015. "Natural Disasters , Household Welfare , and Resilience : Evidence from Rural Vietnam To cite this version : HAL Id : halshs-01109452". *World Development* 70:59–77.

Banerjee, Lopamudra. 2010. "Effects of Flood on Agricultural Productivity in Bangladesh". *Oxford Development Studies* 38, no. 3 (): 339–356. ISSN: 1360-0818. doi:10.1080/13600818.2010.505681. http://www.tandfonline.com/doi/abs/10.1080/13600818.2010.505681.

Bernard, H., et al. 1984. "The Problem of Informant Accuracy: The Validity of Retrospective Data". *Annual Review of Anthropology* 13 (): 495–517. ISSN: 00846570. doi:10.1146/annurev.anthro.13.1.495. http://anthro.annualreviews.org/cgi/doi/10.1146/annurev.anthro.13.1.495.

Bertinelli, Luisito, and Eric Strobl. 2013. "Quantifying the local economic growth impact of hurricane strikes: An analysis from outer space for the caribbean" [**inlang**en]. *Journal of Applied Meteorology and Climatology* 52, no. 8 (): 1688–1697. ISSN: 15588424. doi:10.1175/JAMC-D-12-0258.1. http://journals.ametsoc.org/doi/abs/10.1175/JAMC-D-12-0258.1.

Brammer, Author H, et al. 1990. "Floods in Bangladesh : Geographical Background to the 1987 and 1988 Floods". 156 (1): 12–22.

Chaney, Eric. 2013. "Revolt on the Nile: Economic Shocks, Religion, and Political Power". *Econometrica* 81 (5): 2033–2053. ISSN: 0012-9682. doi:10.3982/ECTA10233. http://doi.wiley.com/10.3982/ECTA10233.

Chantarat, Sommarat, et al. 2016. "Floods and Farmers: Evidence from the Field in Thailand". https://econpapers.repec.org/scripts/redir.pf?u=https%3A%2F%2Fwww.pier.or.th%2Fwp-content%2Fuploads%2F2016%2F08%2Fpier_dp_040.pdf;h=repec:pui:dpaper:40..

Chen, Joyce J., et al. 2017. "Validating Migration Responses to Flooding Using Satellite and Vital Registration Data". *American Economic Review* 107, no. 5 (): 441–445. ISSN: 0002-8282. doi:10.1257/aer.p20171052. http://pubs.aeaweb.org/doi/10.1257/aer.p20171052.

Christensen, Darin, and Thiemo Fetzer. *Using R To Estimate Spatial HAC Errors Per Conley.* http://www.trfetzer.com/using-r-to-estimate-spatial-hac-errors-per-conley/.

Colin Cameron, A., and Douglas L. Miller. 2015. "A Practitioner's Guide to Cluster-Robust Inference". *Journal of Human Resources* 50, no. 2 (): 317–372. ISSN: 0022-166X. doi:10.3368/jhr.50.2.317. http://jhr.uwpress.org/lookup/doi/10.3368/jhr.50.2.317.

Conley, T.G. 1999. "GMM estimation with cross sectional dependence". *Journal of Econometrics* 92, no. 1 (): 1–45. ISSN: 0304-4076. doi:10.1016/S0304-4076(98)00084-0. https://www.sciencedirect.com/science/article/pii/S0304407698000840.

Dell, Melissa, et al. 2012. "Temperature Shocks and Economic Growth: Evidence from the Last Half Century". *American Economic Journal: Macroeconomics* 4, no. 3 (): 66–95. ISSN: 1945-7707. doi:10.1257/mac.4.3.66. http://pubs.aeaweb.org/doi/10.1257/mac.4.3.66.

Donaldson, Dave, and Adam Storeygard. 2016. "The View from Above: Applications of Satellite Data in Economics". *Journal of Economic Perspectives* 30 (4): 171–198. ISSN: 0895-3309. doi:10.1257/jep.30.4.171.

Elliott, Robert J.R., et al. 2015. "The local impact of typhoons on economic activity in China: A view from outer space" [**inlang**en]. *Journal of Urban Economics* 88 (C): 50–66. ISSN: 00941190. doi:10.1016/j.jue.2015.05.001. https://ideas.repec.org/a/eee/juecon/v88y2015icp50-66.html.

Felkner, John, et al. 2009. "Impact of Climate Change on Rice Production in Thailand". *American Economic Review* 99 (2): 205–210. ISSN: 0002-8282. doi:10.1257/aer.99.2.205.

Gao, Bo-cai. 1996. "NDWI—A normalized difference water index for". *Remote Sensing of Environment* 266 (April): 257–266. ISSN: 00344257. doi:10.1016/S0034-4257(96)00067-3.

Garbero, Alessandra, and Raya Muttarak. 2013. "Impacts of the 2010 droughts and floods on community welfare in rural Thailand: Differential effects of village educational attainment" [**inlang**en]. *Ecology and Society* 18, no. 4 (). ISSN: 17083087. doi:10.5751/ES-05871-180427. http://pure.iiasa.ac.at/id/eprint/10411/.

Germanwatch. 2018. "Globaler Klima-Risiko-Index 2018", no. Cop 23. https://germanwatch.org/en/download/20396.pdf.

Gröger, André, and Yanos Zylberberg. 2016. "Internal labor migration as a shock coping strategy: Evidence from a typhoon". *American Economic Journal: Applied Economics* 8, no. 2 (): 123–153. ISSN: 19457790. doi:10.1257/app.20140362. http://pubs.aeaweb.org/doi/10.1257/app.20140362.

Guha–Sapir, D, et al. 2012. "Annual Disaster Statistical Review 2010: The Numbers and Trends ; 2011". *Accessed on.* http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:Annual+Disaster+Statistical+Review+2011+The+numbers+and+trends#7.

Guiteras, Raymond, et al. 2015. "Satellites, Self-reports, and Submersion: Exposure to Floods in Bangladesh". *American Economic Review* 105, no. 5 (): 232–236. ISSN: 0002-8282. doi:10.1257/aer.p20151095. http://pubs.aeaweb.org/doi/10.1257/aer.p20151095.

IPCC. 2014. *Climate change 2014. Synthesis report. Versión inglés*, 2–26. ISBN: 9789291691432. doi:10.1017/CBO9781107415324.

Islam, Mir Rabiul, et al. 2018. "From coping to adaptation: Flooding and the role of local knowledge in Bangladesh". *International Journal of Disaster Risk Reduction* 28 (): 531–538. ISSN: 22124209. doi:10.1016/j.ijdrr.2017.12.017.

Kocornik-Mina, Adriana, et al. 2016. "Flooded Cities". https://econpapers.repec.org/paper/cepcepcnp/463.htm.

Lertamphainont, Sirikarn, and Robert Sparrow. 2016. "The Economic Impacts of Extreme Rainfall Events on Farming Households: Evidence from Thailand". https://econpapers.repec.org/scripts/redir.pf?u=https%3A%2F%2Fwww.pier.or.th%2Fwp-content%2Fuploads%2F2016%2F10%2Fpier_dp_045.pdf;h=repec:pui:dpaper:45.

Leys, Christophe, et al. 2013. "Detecting outliers: Do not use standard deviation around the mean, use absolute deviation around the median". *Journal of Experimental Social Psychology* 49, no. 4 (): 764–766. ISSN: 0022-1031. doi:10.1016/J.JESP.2013.03.013. https://www.sciencedirect.com/science/article/pii/S0022103113000668.

Marks, Danny. 2011. "Climate Change and Thailand: Impact and Response". *Contemporary Southeast Asia* 33 (2): 229–258. ISSN: 0129797X. http://www.jstor.org/stable/41288828.

McCarthy, Nancy, et al. 2018. "Shelter from the Storm? Household-Level Impacts of, and Responses to, the 2015 Floods in Malawi". *Economics of Disasters and Climate Change* 2 (3): 237–258. ISSN: 2511-1280. doi:10.1007/s41885-018-0030-9.

Miller, Jeff. 1991. "Short Report: Reaction Time Analysis with Outlier Exclusion: Bias Varies with Sample Size". *The Quarterly Journal of Experimental Psychology Section A* 43 (4): 907–912. doi:10.1080/14640749108400962. eprint: https://doi.org/10.1080/14640749108400962. https://doi.org/10.1080/14640749108400962.

Mohan, Preeya, and Eric Strobl. 2017. "The short-term economic impact of tropical Cyclone Pam: an analysis using VIIRS nightlight satellite imagery". *International Journal of Remote Sensing* 38, no. 21 (): 5992–6006. ISSN: 0143-1161. doi:10.1080/01431161.2017.1323288. https://www.tandfonline.com/doi/full/10.1080/01431161.2017.1323288.

Newey, Whitney K., and Kenneth D. West. 1987. "A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix". *Econometrica* 55, no. 3 (): 703. ISSN: 00129682. doi:10.2307/1913610. https://www.jstor.org/stable/1913610?origin=crossref.

Noy, Ilan, et al. 2019. "Floods and Spillovers: Households After the 2011 Great Flood in Thailand". *Economic Development and Cultural Change* (). ISSN: 0013-0079. doi:10.1086/703098. https://www.journals.uchicago.edu/doi/10.1086/703098.

Paxson, Christina H. 1992. "American Economic Association Using Weather Variability to Estimate the Response of Savings to Transitory Income in Thailand". *The American Economic Review* 82 (1): 15–33.

Poapongsakorn, Nipon, and Pitsom Meethom. 2012. *Impact of the 2011 Floods, and Flood Management in Thailand.* Tech. rep. December. Jakarta: ERIA Research Project Report 2011-8.

Roggemann, Hanne. 2015. "Empirical Essays on Consumption Smoothing : How Rural Households in Vietnam Insure against Income Shocks". PhD thesis.

Rummukainen, Markku. 2012. "Changes in climate and weather extremes in the 21st century". *Wiley Interdisciplinary Reviews: Climate Change* 3 (2): 115–129. ISSN: 1757-7799, visited on 02/19/2019. doi:10.1002/wcc.160. https://onlinelibrary.wiley.com/doi/abs/10.1002/wcc.160.

Sakamoto, Toshihiro, et al. 2013. "Agro-ecological Interpretation of Rice Cropping Systems in Flood-prone Areas using MODIS Imagery". *Photogrammetric Engineering & Remote Sensing* 75 (4): 413–424. ISSN: 00991112. doi:10.14358/pers.75.4.413.

Sayama, T., et al. 2017. "An emergency response-type rainfall-runoff-inundation simulation for 2011 Thailand floods". *Journal of Flood Risk Management* 10 (1): 65–78. ISSN: 1753318X. doi:10.1111/jfr3.12147.

Sirisankanan, Aeggarchat. 2016. "Do agricultural households share risks in Thailand? Evidence from Thai socio-economic panel survey data". *Asia-Pacific Development Journal* 20 (2): 89–108. doi:10.18356/eed20962-en.

United States Department of Agriculture. *South East Asia - Crop Calendars.* https://ipad.fas.usda.gov/rssiws/al/crop_calendar/seasia.aspx.

Xu, Hanqiu. 2006. "Modification of normalised difference water index (NDWI) to enhance open water features in remotely sensed imagery". *International Journal of Remote Sensing* 27 (14): 3025–3033. ISSN: 13665901. doi:10.1080/01431160600589179.

Yang, Dean, and Hwa Jung Choi. 2007. "Are remittances insurance? Evidence from rainfall shocks in the Philippines". *World Bank Economic Review* 21 (2): 219–248. ISSN: 02586770. doi:10.1093/wber/lhm003.

A Processing of flood map data

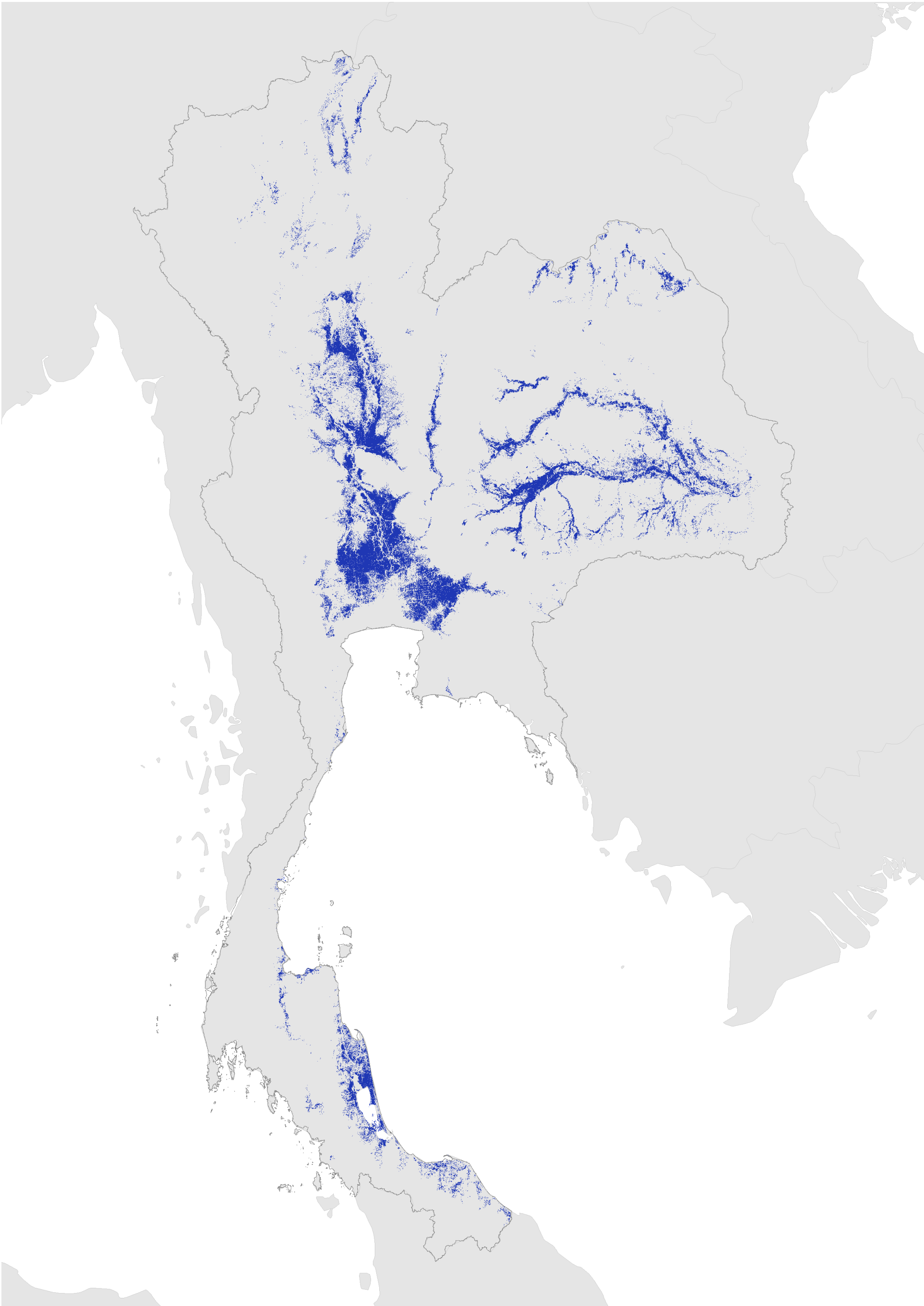


Figure A.1: GISTDA annual map for flooding in Thailand in 2010

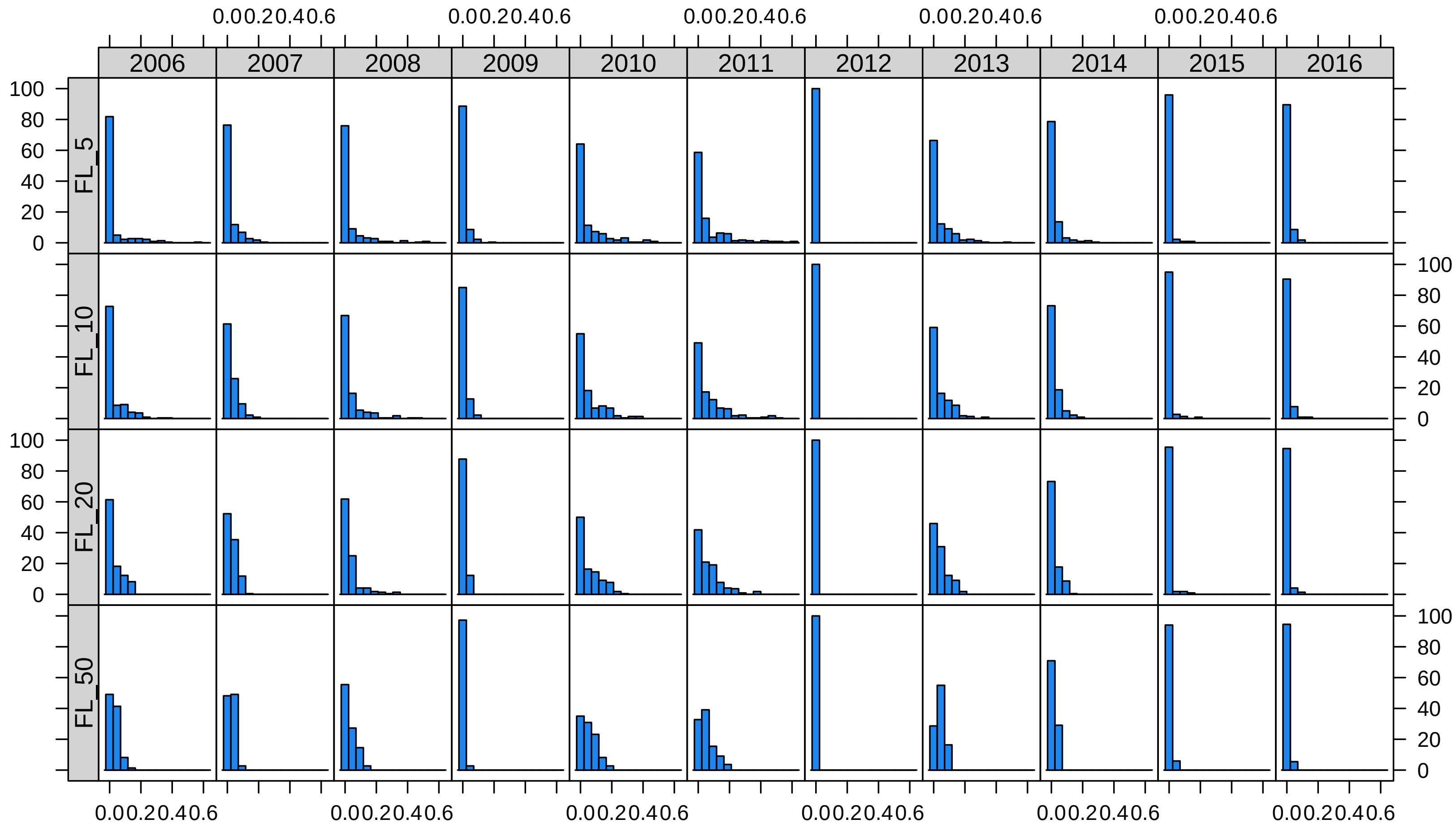


Figure A.2: Histograms showing the distribution of the FL indicator for all villages for 5, 10, 20 and 50 km radii and all survey years. Most of the time villages experience no or very little deviation from the median flood exposure, indicated by a FD taking the value 0. Note the years 2010 and 2011, which were devastating flood years which is also reflected in our data, showing large positive deviations from the median flood exposure. Equally note the years 2012 and 2015, which were drought years. Nearly no village shows a positive deviation from the median flood exposure in this years.

B Processing of TVSEP data

B.1 Data Cleaning

Most of the necessary data cleaning was already carried out by the Thai Vietnam Socio-Economic Survey (TVSEP) team. Nonetheless, some observations that seemed unreasonable or could potentially drive our results still remained in the dataset. A large share of such variables are related to income or expenditure variables. There are two potential reasons for such observations namely outliers or measurement error. While the former is caused by extreme values within a sample, the latter is due to either errors in the collection of data or false recall of individuals when retrospective data is collected. Such issues are more broadly described called informant accuracy problems as (Bernard et al. 1984).

In order to get rid of outliers, there are many possible strategies. The one that seemed most promising to us is the approach suggested by Leys et al. (2013). According to Leys et al. (2013), to detect outliers, one should use the absolute deviation around the median rather than the absolute deviation around the mean. Taking the deviation from the median rather than the mean solves a variety of issues related to distributional assumptions of the data at hand, sensitivity to outliers and extreme values as well as small sample properties (Leys et al. 2013). However, applying this approach to our data lead to dropping around 17% of our observations and had little to no effect on the tails of the distribution of our income and expenditure variables. This lead us to believe that the unreasonable observations in our sample are caused by informant accuracy issues.

To take care of unreasonable observations caused by informant accuracy issues, we take an approach aimed at restricting the standard deviation we allow within a household over the 6 waves. To do so, we drop all the observations whose standard deviation lie beyond 2.5 times of the standard deviation of the respective variable over the 6 years. This threshold level of the standard deviation is in line with suggestions of (Miller 1991). As a result, we drop only 201 out of 11,083 observations for the income per household member variable compared to the previous method in which 1,244 were dropped. In a next step, we apply the same criteria for the different expenditure categories. The only exception to this rule are the Food and Non Food categories for which we are more conservative and only drop observations which lie beyond 4.5 times the standard deviation. This is due to the fact that the standard deviation of these two expenditure categories is smaller than the one of other expenditure variable. Consequently using a threshold of 2.5 would result in dropping a large amount of observation. Moreover it would change the distribution of these two variables in a way that seems unreasonable to us.

By using this method and dropping only 881 observations (7.8% of the sample), both income and expenditures variables seem more evenly distributed making it less likely that our results are driven by extreme observations.

B.2 Creation of Variables

- **Education, literacy, and numeracy variables:** The TVSEP dataset includes detailed information about literacy, numeracy and educational attainment of individual household members. To run allow an easy to interpret analysis of such variables, we only considered the respective characteristics of the household head.
- **Income Variables:** The aggregated income variables are computed manually by us since the there was no information available on the aggregation process of the provided aggregate variables. By adding up all the different income sources we are able to compute an aggregate income variable.

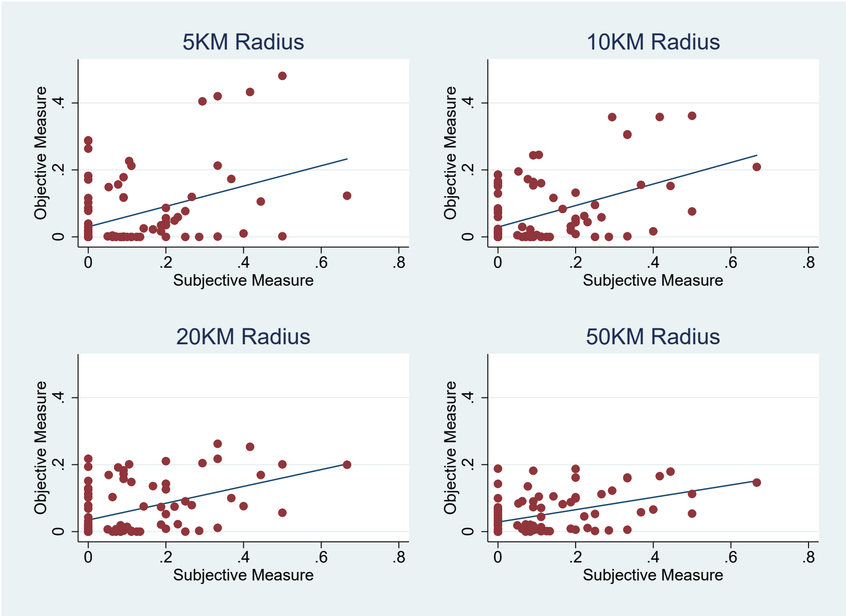


Figure C.1: Subjective VS Objective measure for 2010

- **Insurance variables:** The aggregated measure for the individual types of insurances take a value of one whenever one of the household members answered that they had a specific type of insurance. The dataset includes information on life, property, health, disability health, livestock, funeral, accident, and complete insurance. Moreover, it also includes information on contributions to a government pension scheme.

- **Savings:** The savings variable were collected on the household level. However, due to listing single savings accounts individually, a large share of households and thus household identifiers showed up in the dataset multiple times. Since the type of savings account or the location of the savings where not of immediate interest for our analysis, we aggregated the value of all savings as well as the income from interest on these savings per household by summing them up. With regard to the distance to the institutions where the savings are kept, we weighted the distance to different locations of savings by the value of the respective savings. The values of this variable are thus a weighted average of distances to (potentially) multiple locations.

- **Expenditure Variables:** The different expenditure categories were collected at an individual level, therefore we first proceeded to get the aggregate expenditure values per household. We then converted all the values into the same time unit, in this case, yearly. Once all the items were aggregated and converted into the same magnitude we created 6 categories: Food, Non-Food, Social, Health, Education and Transport & Communication, in order to evaluate potential differences in the behaviour of these 6 groups. With regards to the total expenditures, this variable has been created manually summing up all categories.

C Generating the Subjective measure and its correlation with the objective

The basis of our subjective measure is to proxy the intensity and severity of shocks by using the survey questions related to shocks. We have proceeded as follows: First we have generated a

Table C.1: Correlation Coefficients

	Percentage Flooded				Deviation from Median			
	F5	**F10**	**F15**	**F50**	**F5**	**F10**	**F15**	**F50**
2006	0.066	0.117	0.114	0.046	0.031	0.043	0.025	0.0348
2007	0.035	0.049	0.063	0.030	0.022	0.025	0.043	0.049
2009	0.229	0.11	0.067	0.080	0.030	0.015	0.0031	0.052
2010	0.404	0.486	0.462	0.480	0.382	0.474	0.480	0.470
2012	0.08	0.041	0.056	0.123	0.207	0.171	0.066	0.025
2015	0.314	0.277	0.288	0.243	0.11	0.004	0.0307	0.107

Notes: This table shows the correlation coefficients of the subjective flood measure, as calculated from survey responses, and the subjective flood measure constructed through flood maps. The first four columns show the percentage of area flooded according GISTDA flood maps. The last four columns shows the deviation from the median level of inundation.

dummy which equals to 1 if the given individual reported that flooding was the worst shock it experienced during the previous year. Then we have calculated the share of households within a village for which any of the members reported that flooding was the most severe shock they faced last year. This share is what we call the subjective measure and it has been used to test the association between the perceived impact of flooding and its objective measure. To test this association we have calculated up to 8 correlation coefficients per wave, corresponding to the 4 radius and the two different measures "Percentage Flooded" and "Deviation from the Median". In this respect, it has not been necessary to check the correlation between the area flooded because it is a linear combination of the percentage flooded, therefore the correlation coefficient would be the same. The results of the 48 correlation coefficients can be found in table C.1.

Running the correlations between the two measures we find that there is a substantial positive association between some of them, while for the rest the correlation is virtually zero. Among the six waves, we find the strongest correlation for 2010, year in which the share of area flooded was the largest. In this case, the coefficients for the correlation between the percentage area flooded and the share of household which reported flooding as the most damaging shock are 0.404, 0.486 and 0.462 for 5km 10km and 20km radius respectively. An illustrative picture of this correlation for the 4 radii of interest is shown in figure C.1.

Although it seems that both measures move in the same direction, the correlation is weak. However, this is aligned with the current literature on the association between actual shocks and how they are self-reported. For instance, Guiteras et al. (2015) show that rainfall and self-reported exposure are weak proxies for true flood exposure. Therefore it is not surprising that our subjective measure is not particularly good at explaining the existence and severity of flooding.

D Additional Regression results

Table D.1: The role of educational attainment on coping with flood events

	Dependent variable:		
	log_phm_income	log_phm_netincome	log_phm_farm_income
	(1)	(2)	(3)
FD below	−0.058	−0.023	−0.099
	(0.040)	(0.043)	(0.067)
FD above	−0.099***	−0.081**	−0.107
	(0.036)	(0.041)	(0.068)
FD extreme	−0.089*	−0.075	−0.206**
	(0.052)	(0.061)	(0.084)
I(1 - maxprim)	0.165**	0.138*	0.049
	(0.083)	(0.082)	(0.115)
FD below:I(1 - maxprim)	−0.021	0.025	0.134
	(0.101)	(0.099)	(0.155)
FD above:I(1 - maxprim)	0.010	0.060	0.127
	(0.097)	(0.101)	(0.171)
FD extreme:I(1 - maxprim)	0.103	0.060	0.126
	(0.145)	(0.142)	(0.195)
Rainfall	Yes	Yes	Yes
Rainfall^2	Yes	Yes	Yes
Observations	7,463	7,396	5,489
R^2	0.488	0.488	0.620
Adjusted R^2	0.325	0.324	0.454
Residual Std. Error	0.821 (df = 5668)	0.872 (df = 5601)	1.180 (df = 3819)

Note: This table shows the effect of the three different levels of flooding on different income measures depending on their level of education, where `I(1-maxprim)` is an indicator that the household head has more than primary education. `log_phm_income` measures income per household member. `log_phm_netincome` measure income per household members excluding remittances and public transfers. `log_phm_farm_income` measures income per household member generated from farming activities including income from both, crops and livestock. Besides controlling for rainfall and squared rainfall, all specifications control for changes in household size and the gender of the household head.
Significance levels are given as follows: *p<0.1; **p<0.05; ***p<0.01

Table D.2: Effects of flooding on the probability to receive remittances and public transfers

	Dependent variable:		
	Received Remittances	Received Public Transf.	Disaster Share in PT
	(1)	(2)	(3)
FDC_5below	0.001	0.030	0.081
	(0.007)	(0.020)	(0.066)
FDC_5high	0.00004	0.012	0.095
	(0.007)	(0.022)	(0.066)
FDC_5extreme high	−0.002	0.035	0.040
	(0.007)	(0.025)	(0.076)
Rainfall	Yes	Yes	Yes
Rainfall^2	Yes	Yes	Yes
Observations	5,022	8,028	889
R^2	0.716	0.605	0.923
Adjusted R^2	0.557	0.491	0.597
Residual Std. Error	0.141 (df = 3225)	0.356 (df = 6226)	0.245 (df = 169)
Note:			*p<0.1; **p<0.05; ***p<0.01

Note: This table shows the effect of the three different levels of flooding on the probability of households receive remittances, public transfers as well as the share of transfers received by households due to exposure to floods. Remittances include all remittances obtained from household members, friends, or relative. Besides controlling for rainfall and squared rainfall, all specifications control for changes in household size, the gender as well as the educational attainment of the household head.

Significance levels are given as follows: *p<0.1; **p<0.05; ***p<0.01

Table D.3: Effects of flooding on various income variables split in female and male household head samples

	Dependent variable:		
	log_phm_income	log_phm_netincome	log_phm_farm_income
	(1)	(2)	(3)
FDC_5below	−0.020	0.016	−0.129*
	(0.044)	(0.046)	(0.071)
FDC_5high	−0.099**	−0.087*	−0.202**
	(0.041)	(0.046)	(0.087)
FDC_5extreme high	−0.106*	−0.093	−0.270***
	(0.056)	(0.065)	(0.095)
fem	0.045	−0.069	−0.167
	(0.059)	(0.069)	(0.145)
FDC_5below:fem	−0.134**	−0.111*	0.073
	(0.063)	(0.063)	(0.138)
FDC_5high:fem	−0.005	0.045	0.279**
	(0.067)	(0.072)	(0.126)
FDC_5extreme high:fem	0.117	0.104	0.142
	(0.078)	(0.080)	(0.148)
Rainfall	Yes	Yes	Yes
Rainfall^2	Yes	Yes	Yes
Observations	7,795	7,729	5,718
R^2	0.481	0.483	0.609
Adjusted R^2	0.326	0.326	0.444
Residual Std. Error	0.820 (df = 5993)	0.873 (df = 5927)	1.200 (df = 4020)

Note: This table shows the effect of the three different levels of flooding on different income measures depending on their level of education, where `fem` is a dummy indicating the the household head is a female. `log_phm_income` measures income per household member. `log_phm_netincome` measure income per household members excluding remittances and public transfers. `log_phm_farm_income` measures income per household member generated from farming activities including income from both, crops and livestock. Besides controlling for rainfall and squared rainfall, all specifications control for changes in household size and the educational attainment of the household head. Significance levels are given as follows: *p<0.1; **p<0.05; ***p<0.01

Table D.4: Effects of flooding on various expenditure variables split in samples of household head with maximum primary education and more than primary education

	Dependent variable:			
	TotExp	NFoodexp	Foodexp	healthexp
	(1)	(2)	(3)	(4)
FDC_5below	−0.063*	−0.035	−0.029	0.038
	(0.032)	(0.024)	(0.033)	(0.083)
FDC_5high	0.007	−0.010	−0.008	−0.035
	(0.038)	(0.020)	(0.035)	(0.088)
FDC_5extreme high	0.013	−0.011	0.058	0.147
	(0.040)	(0.030)	(0.037)	(0.105)
I(1 - maxprim)	0.002	−0.013	0.079	−0.127
	(0.051)	(0.046)	(0.060)	(0.176)
FDC_5below:I(1 - maxprim)	0.101	0.011	0.144*	0.119
	(0.074)	(0.057)	(0.077)	(0.271)
FDC_5high:I(1 - maxprim)	0.134	−0.001	−0.019	0.201
	(0.084)	(0.068)	(0.074)	(0.212)
FDC_5extreme high:I(1 - maxprim)	0.184*	−0.007	0.111	0.002
	(0.099)	(0.099)	(0.103)	(0.274)
Rainfall	Yes	Yes	Yes	
Rainfall^2	Yes	Yes	Yes	
Observations	7,661	7,331	6,994	4,807
R^2	0.578	0.608	0.502	0.502
Adjusted R^2	0.449	0.481	0.331	0.228
Residual Std. Error	0.643 (df = 5861)	0.534 (df = 5532)	0.645 (df = 5198)	1.300 (df = 3099)

Note: This table shows the effect of the three different levels of flooding on different expenditure measures depending on their level of education, where `I(1-maxprim)` is an indicator that the household head has more than primary education. `TotExp` measures log total expenditure per household member. `NFoodexp` measures log non-food expenditure per household members. `Foodexp` measures log food expenditure per household member. `healthexp` measures log health expenditures per household member. Besides controlling for rainfall and squared rainfall, all specifications control for changes in household size and the gender of the household head.
Significance levels are given as follows: *p<0.1; **p<0.05; ***p<0.01

Table D.5: Effects of flooding on various expenditure variables split in female and male household head samples

	Dependent variable:			
	TotExp	NFoodexp	Foodexp	healthexp
	(1)	(2)	(3)	(4)
FDC_5below	−0.019	−0.018	−0.018	0.061
	(0.038)	(0.029)	(0.035)	(0.098)
FDC_5high	0.041	−0.024	−0.020	−0.014
	(0.041)	(0.026)	(0.033)	(0.091)
FDC_5extreme high	0.031	−0.045	0.065	0.181*
	(0.038)	(0.037)	(0.041)	(0.100)
fem	0.101**	0.085*	0.021	0.193
	(0.050)	(0.046)	(0.042)	(0.132)
FDC_5below:fem	−0.080	−0.064	0.002	−0.136
	(0.056)	(0.042)	(0.045)	(0.145)
FDC_5high:fem	−0.053	0.035	0.027	−0.030
	(0.067)	(0.050)	(0.054)	(0.120)
FDC_5extreme high:fem	0.030	0.087	0.021	−0.085
	(0.060)	(0.058)	(0.062)	(0.151)
Rainfall	Yes	Yes	Yes	
Rainfall^2	Yes	Yes	Yes	
Observations	7,996	7,654	7,305	4,996
R^2	0.573	0.605	0.499	0.502
Adjusted R^2	0.449	0.484	0.336	0.236
Residual Std. Error	0.647 (df = 6191)	0.535 (df = 5850)	0.644 (df = 5503)	1.295 (df = 3257)

Note: This table shows the effect of the three different levels of flooding on different expenditure measures depending on their level of education, where `fem` is an indicator that the household head is female. `TotExp` measures log total expenditure per household member. `NFoodexp` measures log non-food expenditure per household members. `Foodexp` measures log food expenditure per household member. `healthexp` measures log health expenditures per household member. Besides controlling for rainfall and squared rainfall, all specifications control for changes in household size and educational attainment of the household head.
Significance levels are given as follows: *p<0.1; **p<0.05; ***p<0.01

Table D.6: Effects of flooding on various income variables using flood indices created with NASA flood data

	Dependent variable:		
	log_phm_income	log_phm_netincome	log_phm_farm_income
	(1)	(2)	(3)
FCD_NASA_A_5extreme low	0.063	0.098	0.018
	(0.051)	(0.062)	(0.110)
FCD_NASA_A_5low	−0.025	−0.014	0.082
	(0.038)	(0.046)	(0.066)
FCD_NASA_A_5high	0.007	0.011	0.007
	(0.032)	(0.035)	(0.063)
FCD_NASA_A_5extreme high	−0.030	−0.032	−0.073
	(0.028)	(0.030)	(0.051)
Observations	8,910	8,829	6,610
R^2	0.479	0.483	0.608
Adjusted R^2	0.327	0.330	0.449
Residual Std. Error	0.811 (df = 6892)	0.868 (df = 6811)	1.190 (df = 4701)

Note: This table shows the effect of the three different levels of flooding on different income measures. `log_phm_income` measures income per household member. `log_phm_netincome` measure income per household members excluding remittances and public transfers. `log_phm_farm_income` measures income per household member generated from farming activities including income from both, crops and livestock. Besides controlling for rainfall and squared rainfall, all specifications control for changes in household size, the gender as well as the educational attainment of the household head.
Significance levels are given as follows: *p<0.1; **p<0.05; ***p<0.01

Table D.7: Effects of flooding on various expenditure variables using flood indices created with NASA flood data

	Dependent variable:			
	TotExp	NFoodexp	Foodexp	healthexp
	(1)	(2)	(3)	(4)
FCD_NASA_A_5extreme low	−0.037	−0.043	−0.049	−0.195*
	(0.043)	(0.032)	(0.044)	(0.102)
FCD_NASA_A_5low	−0.034	−0.011	−0.063**	−0.081
	(0.024)	(0.020)	(0.028)	(0.065)
FCD_NASA_A_5high	0.030	0.024	−0.032	0.030
	(0.023)	(0.017)	(0.021)	(0.060)
FCD_NASA_A_5extreme high	−0.048**	0.023	−0.026	0.064
	(0.023)	(0.020)	(0.027)	(0.054)
Rainfall	Yes	Yes	Yes	
Rainfall^2	Yes	Yes	Yes	
Observations	9,125	8,753	8,327	5,764
R^2	0.574	0.604	0.510	0.497
Adjusted R^2	0.453	0.486	0.353	0.240
Residual Std. Error	0.647 (df = 7103)	0.537 (df = 6732)	0.640 (df = 6308)	1.299 (df = 3814)

Note: This table shows the effect of the three different levels of flooding on different expenditure measures. `TotExp` measures log total expenditure per household member. `NFoodexp` measures log non-food expenditure per household members. `Foodexp` measures log food expenditure per household member. `healthexp` measures log health expenditures per household member. Besides controlling for rainfall and squared rainfall, all specifications control for changes in household size, the gender as well as the educational attainment of the household head.
Significance levels are given as follows:*p<0.1; **p<0.05; ***p<0.01