Application of Geographic Information Systems (GIS) for Measuring the Impact Income Vulnerability on Rural Households: A Case Study of the 2010 Floods in Pakistan

Umer Saleem[1], Takeshi Mizunoya[2], Yabar Helmut[2], Muhammad Saad Moeen[3] & Ammara Ajmal[4]

[1] Graduate School of Life and Environmental Science, University of Tsukuba, Ibaraki, Japan

[2] Faculty School of Life and Environmental Science, University of Tsukuba, Ibaraki, Japan

[3] International Food Policy Research Institute (IFPRI), Pakistan

[4] Graduate School of Human Science, University of Tsukuba, Ibaraki, Japan

Correspondence: Umer Saleem, Graduate School of Life and Environmental Science, University of Tsukuba, Ibaraki, 1-1-1 Tennodai, Tsukuba, Ibaraki 305-8572, Japan. Tel: 81-29-853-4958. E-mail: sparks.usar@gmail.com; s1736037@u.tsukuba.ac.jp

Abstract

While the economic impact of natural disasters has been studied extensively, there are rather few studies that have addressed their impact on household income. This research tries to fill this gap by analyzing their actual effect on household income caused by the 2010 floods. We look at the impact of natural disasters on rural households in Pakistan after a massive flooding event in 2010. We used the difference-in-difference (DID) approach, which showed statistical significance at 1 percent. We also calculated the household distance from the rivers to see how vulnerable household income is to such kinds of shock-like floods. The results showed that the sample households living near had more impact as compared to the ones living far. Stata and Geographical Information System (GIS) software packages have been used for regression analysis and calculation of distance, respectively. This research will not only give insight in order to understand household income losses but will allow government, policymakers, and International Aid agencies to plan, make countermeasure strategies before designing post-disaster projects. After taking into account the effect of floods, which tend to have far more impact on the households, which are located near the source of the flooding. In this case, they need to focus more in terms of reconstruction of infrastructure, particularly for the households which are near these flooded areas. Firstly, this finding can give policymakers insight in terms of strategies to develop agriculture and non-agriculture employment opportunities. Secondly, it is essential to reduce income vulnerability and improve rural household finance economic conditions.

Keywords: rural, household, flood, income, natural disasters, vulnerability, difference-n-difference (treatment effect model)

1. Introduction

The occurrences of flooding in Asia are common, like in China, India, and Bangladesh, but Pakistan has also been identified as one of the most vulnerable countries to climate risks and broader hazards in Asia (Kreft et al., 2016). Economic activities in rural areas, especially for developing countries, are mostly dependent on climate as their prime activity is agriculture. Rain, drought, or any change in climatic factors affect rural households income. Vulnerability is defenseless, insecurity, and exposure to events like natural disasters (Chambers, 1989). Households trying to protect their produce from such kinds of events. Economists have studied about these disasters in order to identify the risks and their mechanisms (Hoddinott and Quisumbing, 2003; Heeks and Ospina, 2019). On the other hand, addressing these dangerous conditions and pressures on households will reduce the vulnerability level (Wisner et al., 2004).

The study took the 2010 flooding in Pakistan as our case study to analyze the extent of losses to rural household's income after these natural disasters. This will help us cope with the potential losses, either temporary or permanent. We also will analyze the extent of income losses based on distance from the nearby rivers

We try to add to the body of knowledge while focusing on Pakistan and trying to examine the household income losses with the change in distance. Methodologically, we have put forward an original identification strategy along with the data set PRHPS round 1 as unexpected it focuses on flooding event 2010, which creates it as a quasi-natural experiment. In this paper, we investigate the impact of floods with regards to distance as to see rural households' conditions and then come up with recommendations. To explore that we have two questions in particular and to understand this scenario, and we tried to fill these gaps in the literature by investigating the following questions: (i) To what extent the household income was affected due to flood as mostly the rural households? (ii) Do flood-affected households living near the river are affected more or unaffected by floods? We use difference and difference analysis (DID) and Emily Oster estimates (Oster, 2019), which will be discussed in further sections.

The empirical analysis suggests that households after the event of flooding tend to move away from activities related to agriculture due to which there is a decline in their income. Therefore, the observed changes in income strategies do not necessarily imply a structural change; instead, they reflect flood-affected household's short-term coping with the harms of disaster. The rest of the paper is structured as follows: Section II provides information about the background flooding event of 2010, discussing the modeling, empirical strategy, and data used in this paper, Section III consists of results in which we described the regression, difference-in-difference (DID) analysis, and Oster estimates and, lastly, section IV is about recommendations and conclusions.

2. Background and Empirical Strategy

Pakistan is prone to events like earthquakes, floods, and landslides, but among all the climatic events, *flooding* is the most recurring (see Table 1). The reason for that is the excessive monsoon rainfall and melting of glaciers. Between 1999 and October 2019, Pakistan experienced a total of 63 major floods, resulting in nearly 7000 deaths and $18.8 billion in economic losses (EM-DAT, 2019). Among these recent events, the 2010 flood was particularly damaging. It was one of the most notable floodings events that happened in the country, which impacted the Indus River basin all across Pakistan devastating all provinces of Khyber Pakhtunkhwa, Sindh, Punjab, and Balochistan. From late July to September 2010, the flood-affected almost a fifth of Pakistan's total land area, over 2000 people lost there lives, 20 million people got affected, and economic damages reached approximately $16 billion (World Bank DNA Report, 2010). The agriculture sector also suffered a lot as unharvested crops, including cotton, sugarcane, rice, and vegetables, which covered more than 2.4 million hectares of land, was washed way, which caused $5 billion damage in the agriculture sector (FAO 2015).

Table 1. Number of floods in Pakistan, last 20 years (1999-2019)

Year	Occurrence	Total deaths	Injured	Affected	Homeless	Total affected	Total damage ('000 US$)
1999	2	34	43		1000	1043	
2001	1	210	179	400000		400179	246000
2002	3	37	10	3000	1000	4010	30
2003	3	266	476	1265767		1266243	
2004	2	5					
2005	5	636	470	7523073	3500	7527043	30000
2006	7	400	525	2000	5600	8125	
2007	6	526	206	2500		2706	327118
2008	3	83	12	290752		290764	103000
2009	3	102	80	75000		75080	
2010	4	2113	2946	20360550		20363496	9500000
2011	1	509	755	5400000		5400755	2500000
2012	3	518	2902	5047662		5050564	2500000
2013	2	268	912	1496870		1497782	1500000
2014	1	255	673	2530000		2530673	2000000
2015	6	367	499	1576991		1577490	1000
2016	7	369	267	3135	7360	10762	2000
2017	2	180	817	62200		63017	110000
2018	1	60					
2019	1	25	22			22	
Total	**63**	**6963**	**11794**	**46039500**	**18460**	**46069754**	**18819148**

Notes: Total flood occurrences in Pakistan from 1999 to 2019, indicating the number of deaths, injured, affected, homelessness and damages.

Source: EM-DAT: The Emergency Events Database - Universite catholique de Louvain (UCL) - CRED, D. Guha-Sapir - www.emdat.be, Brussels, Belgium *(Accessed on: June 22, 2019)*

The 2010 floods in Pakistan were caused by hefty monsoon rain in the last days of July and the first few days of August. Although much of the shower, concentrated in the north and north-western parts of Pakistan, places as far south as Mirpur Khas in Sindh received large amounts of rain as well. Areas in the north and north-east that received heavy rains were mostly arid with steep slopes and little vegetation. Which resulted in very rapid runoff and caused flash floods. The Kabul, Swat, Sibi, and other smaller rivers in Khyber Pakhtunkhwa and Balochistan, spread outside their banks and caused considerable damage, particularly in districts such as Nowshera, Peshawar, Swat, Jafferabad. Muzaffargarh, D.G Khan, Rajanpur, Dadu, and Kashmore. Many of the main canals that come off the Indus Khyber Pakhtunkhwa, Punjab, Sindh, and Balochistan, some of which had already had heavy rain, the significant rivers filled up.

The Indus, in particular, began to expand into its immediate flood plain and subsequently overflowed its banks and flooded the surrounding areas leading to colossal devastation. As the flood continued into lower Punjab and Sindh, where slopes are minimal, the velocity of water slowed, and the flood spread over a large area. In some areas, the river reached a breadth of 50 km and beyond. The situation in rural areas further aggravated because the flood water got diverted to rural areas to protect major towns. These diversions resulted in some significant canals and dams to overflow, and as a result, water diverted on to agricultural lands. The water from various northern districts entered the open areas of by early to mid-September, the Indus in its upper reaches had begun to recede, but there had been significant changes in its course. However, due to the slow velocity of the water and continuing rain in the north, many parts of Sindh continued to receive floodwater, which spread in many highly productive canal irrigated areas (Figure 2).
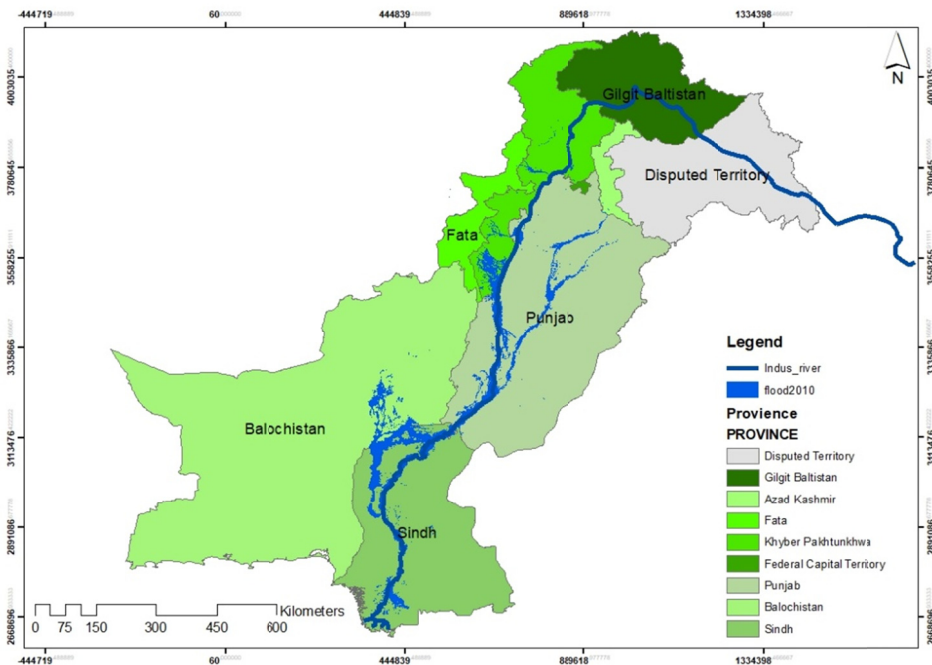
Figure 1. Flood 2010 outflow

Source: RHPS and Pakistan Administrative data from Geofabrik

Moreover, floodwaters remain stagnant in many areas in Balochistan, Punjab, and Sindh. Some of these areas are low lying depressions, but in other areas, roads and railway embankments have also been acting as bunds and limiting drainage. As a result, water is likely to recede for several months due to evaporation and percolation. Rain is expected to absorb more slowly in the heavy soils in these areas that are flood-affected, 111 districts of Pakistan, covering more than 20 percent of the country's territory. Figure 3 shows the flood-affected provinces and districts at the end of September 2010.
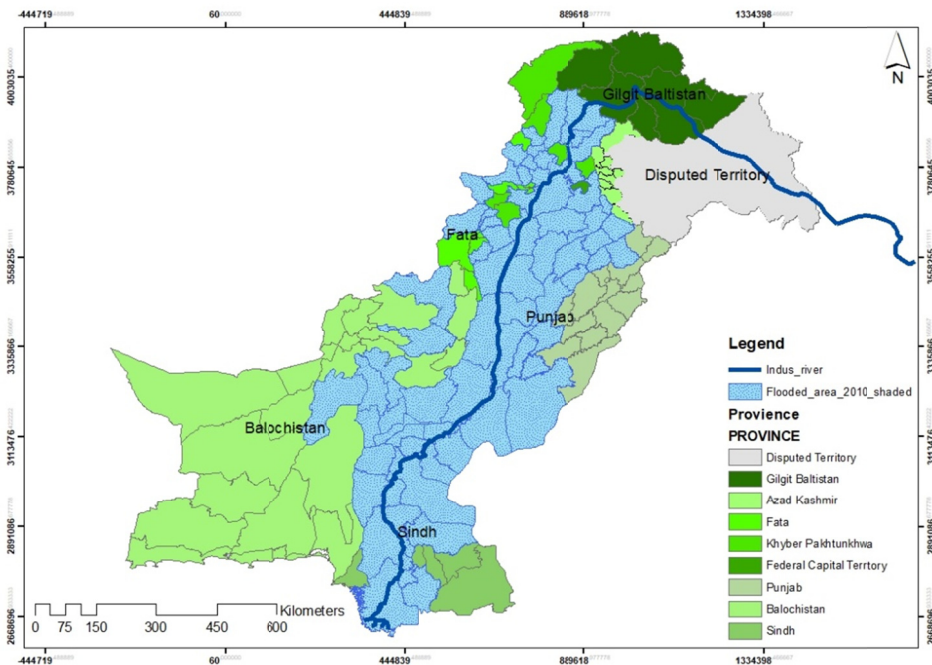

Figure 2. Flood 2010 affected districts by area

Source: RHPS and Pakistan Administrative data from Geofabrik

Economic research on vulnerability and income has been studied extensively on different issues but focusing on quantifying the risks appropriately (Hoddinott and Quisumbing, 2003). Some authors have focused on the conceptual framework (Dercon, 2006; Brigulio, 2008), and others in generating empirical measurements using data from developing countries (Chaudhuri, Jalan, and Suryahadi, 2002; Kamanou and Morduch, 2002; Ligon and Schecter, 2003; Ward, 2016). Recent developments have started adopting these methodologies using more disaggregated approaches, such as considering sub-national and spatially dependent indexes (Naude, McGillivray and Rossouw 2009; Webber and Roussow, 2013). Understanding individual and aggregate vulnerability differences (Calvo, 2018), and testing the accuracy of measurements using panel datasets (Zhang and Wan, 2009; Dutta, Foster and Mishra, 2011; Celidoni, 2012). Other authors explore these measurements further, using different perspectives around the same concept. For instance, defining vulnerability as the insight to poverty status in at least one period in the future and for forecasting trend (Pritchett, Suryahadi and Sumarto 2000; Feeny & McDonald, 2016), by considering risk profiles using expected low utility (Ligon and Schecter2003), and by describing the dynamic nature of this concept regarding future macroeconomic shocks (Glewwe and Hall 1998; Alfani et al., 2019). The choice to measure is essential when examing the ex-post poverty, but it is not as important when analyzing the ex-ante vulnerability with respect to poverty (Azeem et al., 2018).

There are countless new social projects and programs realized worldwide every year. However, most social programs face many challenges. The first problem is that they do not last beyond the initial stage, wasting resources and the hopes placed in them. Secondly, developing countries, which have scares resources, waste most of the financial and human investments put into the projects.

In order to evaluate the actual effect of the flood 2010, we have used the difference-in-difference (DID) approach. As we based our case study on the natural experiment, the DID approach is the most suitable method to adopt with the study design.

2.1 Background about Difference and Difference Method

The DID estimation method is an econometric modeling tool that is well established, though with some issues which we tried to address in this research like omitted variable bias. However, the main components of this approach are well established. Snow in (1854) used the DID approach for the first time in a scientific study for survey-based research. Additionally, Qiu and He in 2017 first applied it in natural sciences. The DID method, while taking into account the impact of loss due to the flood, we have used the control variables along with district fixed effect to address the unbalanced panel problem (Chaudhuri 2003; Lechner, Rodriguez-Planas, and Fernández Kranz 2016; Epping-Jordan et al. 2015; Oster 2019;)

2.2 Data and Variables

The paper utilized data from the Rural Household Panel Survey (RHPS), Round 1.0, conducted in 2012 by the International Food Policy Research Institute, Pakistan (IFPRI), and Innovative Development Strategies (IDS). After cleaning the data, the sample size is 1,985 households, with 1,156 in Punjab, 483 in Sindh, and 77 in Khyber Pakhtunkhwa (KP) province. They applied the multistage sampling methodology for the data collection. The Rural Household Panel Survey (RHPS) data consist of 19 districts of three provinces, which are 12 from Punjab, followed by five from Sindh and two from KP. The total number of villages is 76 in the sample based on four villages from each district. Each village has data of 28 households; therefore, they chose a total of 2,124 households for the survey (Nazli & Haider, 2012).

We analyzed important variables such as gender of household head, household head education level, land ownership of household, annual household expenditure, losses due to flood, household assets, and drainage system. The study calculated household income based on different sources like agriculture and non-agriculture income. The agriculture income consists of income after the sale of agriculture produce and its byproducts. After that, we subtracted all the expenditure that occurred during the production like primary labor, material, equipment, pesticides, seeds, transportation, and others. Whereas for non-agriculture income consist of non-farm business, remittances, other sources of income. As for other sources of income includes household which either gets a pension, rent from the property which comprises of their income. After that, both the agriculture and non-agriculture income incomes were added to get the household's income. We also computed the rainfall and distance of the household from the flooded area. The amount of rain received during 2010 is calculated for the six-month average and then added them to see the total rainfall received during the monsoon season (from June to November measured in mm/hr.). Likewise, the elevation of the households is also taken into account and calculated with the help of ArcGIS. We made these two variables in quantiles to see the correlation in different ranges. Lastly, the distance of each household was calculated based on the household location data from RHPS

from the ArcGIS open source database. The data based on the Global Positioning System (GPS) for the household location and administrative data for Pakistan taken from DIVA-GIS and Geofabrik. These datasets import in the ArcGIS software to verify that the household location is given correctly in the data (Appendix Figures 1, 2, and 3).

The data consist of 13,376 members in 1984 households, of which include the average mean household age is 46.32, but it consists of all age groups. Fifty-one percent of the household surveyed were male, and 49 percent are female respondents; 49 percent of the household went to school. If we see the division in terms of occupations, 49 percent of the households are the agricultural households, and livestock households which own livestock are 68 percent. Lastly, about 10 percent of the household reported they were being affected by the flood in 2010, accounted for a loss of about Rs. 43,083, on average, other summary statistics are given below in table 2.

Table 2. Summary statistics

Variables	Obs.	Mean	Std. Dev.
Age of household head (years)	1,984	46.32	13.65
Gender of household head (1=Male and 2= Female)	1,984	1.02	0.13
The education level of household head (1=any level of schooling)	1,984	0.49	0.50
Land ownership of household (Acres)	1,984	3.34	9.60
Household value of agricultural assets as of 2010 (PKR)	1,984	34,016	146,397
Household value of other assets as of 2010 (PKR)	1,984	257,275	360,340
Household value of properties as of 2010 (PKR)	1,984	211,958	315,713
Household savings as of 2010 (PKR)	1,984	3,895	41,056
Household total value as of 2010 (PKR)	1,984	507,144	723,895
Household annual expenditure in 2010 (PKR)	1,984	47,621	52,178
Household agricultural income in 2010 (PKR)	1,984	-30,965	1,146,557
Household non-farm income in 2010 (PKR)	1,984	97,824	182,532
Household total income in 2010 (PKR)	1,984	66,859	1,134,128
Drainage system available at the household (1=yes)	1,984	2.36	0.75
Household affected by flood in 2010 (1=yes)	1,984	0.10	0.30
Total losses of the household due to flood 2010 (PKR)	1,984	43,082	165,307
Total household losses (PKR)	1,984	119,503	244,830
The distance of Household from the flooded area (Km)	1,984	30.84	42.26
Household loss (Log)	1,984	8.94	4.39

Source: The author calculated based on IFPRI- Rural Household Panel Survey Round 1

2.3 Modeling and Empirical Specification

Some authors have explored the concept of vulnerability to natural disasters using sociological perspectives (Fothergilland Peek 2004), socio-demographic factors (Finch, Emrich and Cutter 2010), and practical aspects of human geographies, such as famines (Watts and Bohle 1993). Hazard and disaster specialists often focus on Vulnerability as sources of risk, and how societies cope with them once they hit (Yong, Qi-Fu and Ling 2001). However, it has not been considered in terms of natural disasters, especially in Pakistan, due to the prevalence of location-specific risks in many contexts

We evaluate whether exposure to floods predicts variations in changes in the sources of income and distance from flood using a difference-in-differences approach. The difference-in-differences, DID, of the effect of treatment on the dependent variable of interest is shown in equation 1.

$$DID = \Delta^{\alpha} - \Delta^{\beta} \qquad (1)$$

where Δ^{α} represents the change in the mean of the outcome variable for the treatment group from the

pre-treatment period to the post-treatment period. Δ^{β} provides a similar measure for the control group. The difference, DID indicates the effect of the treatment relative to the control.

We explore whether variations in distance from flooding sources have some significance in terms of income change among rural households from Pakistan. We first hypothesize that households may exhibit accelerated movement from the farm to the non-farm sector for income in response to flood exposure. To explore this, we estimate the difference-in-differences model in equation (1) employing the following random effect model that captures the structural change for household i in time t due to flood exposure:

In the literature, we studied many perspectives; this work will adopt a theoretical definition that is consistent with the idea of the rural household as the probability of falling income after natural disasters. We will use difference -in- difference approach (DID), so Let Y_{isd} be the income measurement for the rural households i, d is the distance and s is the flood-affected area. The following expression equation 2 is obtained.

$$Y_{isd} \;=\; \alpha + \beta_1 D_d + \beta_2 D_s + \beta_3 D_d \times D_s + \beta_4 X_{isd} + \mu_{isd} \qquad (2)$$

$$d = \begin{matrix} 1 & Household\ distance\ from\ flood\ area\ if\ (km) \\ 0 & otherwise \end{matrix}$$

$$s = \begin{matrix} 1 & Treatment\ group:\ Flood\ affected\ household \\ 0 & Control\ group:\ Non-Flood\ affected\ households \end{matrix}$$

Where Yi represents household (income), Please note that for simplicity, we ignore covariate of X_{ist}. β_3 as the main coefficient to estimate the effect of the difference-in-difference estimates, while X_{isd} is a vector of the control variables in equation 2.

It is essential to understand first that the DID method can be applied when two assumptions are satisfied. Firstly, one is a parallel trend assumption that the trends of loss of income over distance should be the same across the experimental and non-experimental households. As with this assumption, the DID ensures to use the control group as the factual countering scenario to the treatment group. In this study, we used the data of distance from 5 km to 60 km from the flooded area, which covers more than 2 points of assessment. Therefore, we can see this assumption through the distance to see the consistency of the preintervention distance trends for the flooded households and other districts.

The main focus of this research is to see the income losses due to disasters (in this case study as flood), which makes our parameter of analysis to be $\beta3$. In our difference-in-differences model, it shows the analysis of the impact of flood over income loss. The crucial assumption for identifying the causal effect of flood the 2010 flood, changes after the shock that the selectivity of people into general and income loss does not vary over time. In other words, we assume that if a decline in income occurs due to flood, changes with the change in the distance from the flood source. We want to see what difference in the income losses will take place with the change of distance. With this approach, we can estimate the impact of flooding in 2010 by separating the households which suffered more losses or not by a change in distance or range.

In the next step, we try to see the common heuristic effect for evaluating robustness for our results. For that, we see omitted variable bias is to observe the movement of the coefficient even after using the control variables. Using Oster estimated, this empirical test allows us to have the information only if selection on observables is informative about selection on the unobservable. The Oster estimates firstly takes into account the coefficient movement and secondly the movements of the value of R-squared, which allow us to identify the omitted variable bias. To prove these estimates, Emily Oster, in 2019, used a broad set of publications in the field of economics and used the evidence from previous randomized studies to draw further guidelines that we used it to address the issue of unobserved variable bias (Oster. 2019).

3. Results

Based on the regression equation (2), we ran different types of regression analysis to see the effect of natural disasters on income. Table 2 shows the results to measure the impact of the flood on household income to show income vulnerability. The satellite images have matched households that were affected by flood during the year 2010 with that of IFPRI household panel survey round 1. The dependent variable is the household loss (taken as log) we combined the all the losses the household had during the year 2010-11 like income loss, agricultural loss,

business loss. The loss of income during the year 2010-11 is due to floods, drought, earthquake, fire, and others (Appendix table 1). The independent variable is the dummy variable if the household is affected by the flood 2010 (1 if yes). As to minimize the variable bias, we have taken the control variable by adding socio-economic, demographic, and community characteristics.

3.1 Ordinary Least Square OLS Estimates

The results begin with the first reporting of OLS estimates based on equation (1), presented in Table 3. Column (1) reports estimates of the correlation between the use of flood-affected households (1=flood affected, otherwise 0) and income loss of household in 2010-11 (taken as log). The coefficient of an estimate is 3.382 and statistically significant at 1 percent p-value. This implies that if the household is affected by the flood, the probability of loss of income increases by 1 percent on average. To check the stability of OLS estimates given in column (1), we used district fixed effect along with control variable for household characteristics such as household head education (any schooling), land ownership of household (Acres), log of annual household expenditure in 2010, primary irrigation of sources (1=yes), drainage system available at the household (1=yes) and quantiles for rainfall in column (2). The coefficient of the main explanatory variable is 3.170 and report statistically significant at 1 percent p-value. Column (3) reports the estimates with control variables for household characteristics and quantiles elevation. The coefficient of the main explanatory variable is 3.569 and report statistically significant at 1 percent p-value. The estimates of column (4) show that the correlation between flood-affected households (1=flood affected, otherwise 0) and income loss of household in 2010-11 for a specification that includes district fixed effects. With the tehsil fixed effect, the correlation between two variables becomes statistically insignificant, which implies that flood-affected households are not being strongly correlated with income loss at the tehsil level.

The results given in column 1-4 show that the households affected by the flood will increase the probability of having a loss of income due to such an event, which is very much understandable as well. These estimates are statistically significant and have an impact on household income with the flood. However, the positive sign implies that household income is having a positive relationship with flooding events in 2010. The OLS estimates are consistent with the theoretical approach of the natural hazard as the households are affected by floods. Whereas, these results are not reliable because of the endogeneity of the choice of energy sources. Therefore, the identification strategy use difference and difference approach (DID) and Emily Oster estimate method to find out the trustworthy results.

In order to understand the OLS estimates, it is essential to understand the statistical significance which we can check by the p-values. The results are statistically significant with and without a robustness check. The mean of the coefficient remains significant after adding control variables. The coefficient also remains stable, and the coefficient does not change after robustness check; it remains within the band, which we discuss in detail later.

Firstly, it is vital to define the economic significance of how big the impact was. In order to see the impact, we look at the standard deviation of the main dependent variable, which household loss, which is 4.39 given in summary statistics table 2. Now to compare it with coefficient from our OLS estimates from our first model in column 1 is 3.38, as in table 4, which is around 77 percent of the standard deviation of the main dependent variable (calculated by 3.38/4.39 = 0.76999). Therefore we can say the impact of flood loss was quite significant because the coefficient is around 77 percent of the standard deviation of the main dependent variable.

As we discussed earlier that the p-values show statistically significant now we explain our results about the OLS estimates based on equation (1), which presented in table 3. Column (1) reports estimates correlation between the flood in 2010 as (1=flooded, otherwise 0) and effect on the income of the household with simple linear regression. The coefficient of an estimate is 3.382 percent and statistically significant at 1 percent p-value. Results imply that if the household that had a decline in its household income in 2010, the main factor which caused it was due to flooding. To check the stability of OLS estimates given in column (1), we used the only district fixed effect, but in further regressions, we have used different control variables. The analysis shows that if the household affected by the flood 2010 has the probability of income vulnerability will increase by 3.382 percent with a change of 1 percent. To check the stability of OLS estimates, we used a district fixed effect in a column, all with other control variables in column (2) (3) (4) like household and community characteristics. Column (2) combined with the household characteristics and elevations quintiles as to see the effect if elevation quintiles. The value of the coefficient is 3.170 and report statistically significant at 1 percent p-value, but no quintile show any statistical significance. In column (3), we can see the flood 2010 effect is combined household characteristics with the rainfall quintiles received during the monsoon season. The value of the coefficient is 3.569 and report statistically significant at 1 percent p-value. The quintiles 4, which is the last quintile with the

maximum value of rainfall, show 0.588 percent statistically significant at 1 percent p-value, show correlation as the main contributor of loss during flood 2010. Lastly, Column (4) reports the estimates after adding all control variables for household characteristics, elevation quintile rainfall quintile, and community characteristics. In order to see the household total income loss due to flood 2010, the coefficient estimate value is 3.440 percent and statistically significant at 1 percent p-value. The OLS results from column 1 to 4 indicates that the coefficient of the main variables remains stable and consistent. The coefficient remains about 80 percent of the standard deviation of the main dependent variable, which indicates that the impact of the flood was economically significant, and households suffered from severe losses.

Our results indicate the small values of R-square firstly, the coefficients' of the main explanatory variable remain significant at 1 percent p-values, which indicates that our regression model has statistically significant explanatory power (Neter, Kutner, & Nachtsheim, 1996; Kutner, 2005). In general, authors especially in social science favor if R-square is above 10%, below 10% may be problematic, but more than 70% or 80% may also be problematic because of multicollinearity. Usually, it is considered that well-specified models should be having a high value of R-square, but in social science, where it is challenging to specify models, low R-square values can often expect. R-square is typically higher because it is easier to specify complete, well-specified models (Heeringa et al., 2017). However, in the social sciences, where it is hard to determine such modes, low R-square values are often expected. Secondly, the data was collected by IFPRI was not specified just for flood-affected households. Instead, it was collected based on a multistage stratified sampling technique (Nazli & Haider, n.d., 2012) so, the data contain the households which are not affected by the households. Thirdly, in our model when explaining R-Square in a proportional explain variation due to independent variables. Usually, it is considered good if R-square is high enough let say more than 0.5, but it is not necessary, because model can have larger R-squared value even if overall model is insignificant and model may/ or may not full fill the necessary assumption of linear regression model such as normality of residual, homoscedasticity, multicollinearity, autocorrelation etc. Lastly, many researchers have low R square (Singleton, 2007; Wallquist et al., 2010) but the interpretations of the significant variables would stay even if low R-squared models.

Table 3. Effect of flood 2010 over loss in household income

VARIABLES	Dependent variable: Log of household losses in 2010-11			
	(1)	(2)	(3)	(4)
Flood 2010 (1=yes)	*3.382****	*3.170****	*3.569****	*3.440****
	(1.139)	*(1.014)*	*(1.143)*	*(1.109)*
Controls:				
District fixed effect	Yes	Yes	Yes	Yes
Quantiles for Elevation		Yes		Yes
Quantiles for Rainfall			Yes	Yes
Household Characteristics		Yes	Yes	Yes
Community Characteristics				Yes
Constant	10.52***	2.327	1.321	2.057
	(0.228)	(2.138)	(1.770)	(2.360)
Observations	1,984	1,984	1,984	1,984
R^2	0.259	0.304	0.299	0.316

Note: The OLS estimates report the correlation between households affected by flood and household income loss. The OLS estimates given in column (1) report a significant correlation between households affected by flood and household income loss with simple regression. The results remain significant in Column 2, 3, and 4 after adding control for district fixed effect along with village and household characteristics.

The author calculated based on IFPRI- Rural Household Panel Survey Round 1

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

The results given in column 1-4 show that the loss to household income increases the probability the household

was affected by the flood. These estimates are statistically significant and have an impact on income due to floods. However, these results become statistically insignificant if adding control for the tehsil fixed effect along with other characteristics. OLS implies that the correlation between flood and household loss is not consistent. Whereas, these results are not reliable because of the endogeneity due to flooding. Therefore, the identification strategy used Oster Estimate approach to find out the reliable results.

The results given in column 1-4 shows that the household affected by flood 2010, their leading cause of the decline in the income was due to the impact of flood in that year. The OLS estimates are consistent with the theoretical approach of the natural hazard and show the decline in household incomes.

3.2 Difference and Difference (DID) Analysis

Now we move to difference and difference analysis. For our analysis, DID is appropriate to estimate the effects of flood in 2010 on the household income (or losses). In this analysis, as we already have the time internal to see its effect as in 2010. In equation (1), we use the distance to see the change with distance by taking two groups as control and treatment, which are flooded households and non-flooded households. The regression equation (1) is used for running 12 different models in which the coefficient beta β_3 the main coefficient, and we try to see with the change in the distance. We have used 12 different distance levels to see the impact of the flood on the household's total income losses. The distance was categorized and set, based on the mean value 31 km of the sampled households from the PRHS round 1. Based on this, we wanted to see the impact on floods with the households and using control variable are kept the same as used in the OLS estimates for consistency. The results are consistent with the OLS estimated, and with the reality as due to the loss in income, the expenditure show statistically significant at 1 percent p-value.

Now we see the intensity of losses for change of distance from the source of the flood. We calculated the distance of the households from the source of the flood, which in this case, the distance from the river. Regression model results show a high and statistically significant correlation between the household losses and due to flood as in table 3. We can see the DID in table 4 (a) and (b) with the baseline results in column 1 as the total loss for the household in the year 2010. Column 1 is the baseline or control group, but from columns 2 to 13, we have the treatment groups with the change in the distance from the mean ranging from below or equal to 5 km to 60 km or above. We then can see the log losses and flood-affected households in terms of the distance dummy, which we have calculated with the help of GIS mapping software by calculating the near distance from the flood 2010. As the origin of a flood is from the river, we calculated the distance of the households from the flood (Appendix figure 3).

All the 13 models show statistically significant at 1 percent p-value. However, the value of the coefficients in column 3 shows the same value for column 1 of the baseline mean that at a distance 55 km household in the sample size did not have any effect or least effect of the distance the contributed to their loss of income. Whereas, in column 13, which shows the result model for 5 km, we see the effect of the loss due to flood to be the highest, which are understandable and proves that the losses to the households will be higher if they are living closer to the river. As we move from the river, the value of coefficients is becoming less and more towards the baseline. It can be argued that the closer you live to the river, the extent of losses would be more, it is true but we are trying to focus as to see the extent of decline of household income decline. In our results, it shows that although the losses are higher for the households living closer to the river, the coefficient of an estimate is *3.908* percent and statistically significant at 1 percent p-value if the household is living at 5 km. The results also imply that if the household that lives closer had a decline in its household income in 2010 for the modeled result in columns 13 to 2, but at 20 km the coefficient of an estimate is *3.376* percent and statistically significant at 1 percent p-value which declined. Secondly, at 5 km to 50 km, we see an increasing and decreasing trend in the household's losses again up to 60 km after that; there was no observable change. We see an increase and decrease in the level of household losses can be due to many factors that are not the focus of research in this paper, which can be for example, poor infrastructure, level of mitigation strategies, and adaption, etc.

DID results estimate the low values of R-square, but the estimated coefficients of our main variable of interest remain significant at 1 percent p-values. Therefore, the results indicate that our regression model has statistically significant explanatory power, and concern on the small value of R-square will remain invalid.

Table 4a. Loss in income with the change in distance (60 km to 35 km)

Variable	Dependent variable: Log of household losses in 2010-11						
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**
Treatment group (=1 if the household distance from the flooded area)	Baseline	HH =< 60 km	HH =< 55 km	HH =< 50 km	HH =< 45 km	HH =< 40 km	HH =< 35 km
Flood in 2010 (1=yes)	*3.440****	*3.447****	*3.440****	*3.432****	*3.483****	*3.525****	*3.453****
	(1.109)	*(1.110)*	*(1.109)*	*(1.114)*	*(1.107)*	*(1.118)*	*(1.122)*
Control:							
District fixed effect	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*
Household Characteristics	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*
Community Characteristics	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*
Observations	1,984	1,984	1,984	1,984	1,984	1,984	1,984
R^2	0.316	0.316	0.317	0.317	0.316	0.317	0.316

The author calculated based on IFPRI- Rural Household Panel Survey Round 1

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

Table 4b. Loss in income with the change in distance (30 km to 05 km)

Variable	Dependent variable: Log of household losses in 2010-11					
	(8)	**(9)**	**(10)**	**(11)**	**(12)**	**(13)**
Treatment group (=1 if household distance from flooded area)	HH =< 30 km	HH =< 25 km	HH =< 20 km	HH =< 15 km	HH =< 10 km	HH =< 05 km
Flood in 2010 (1=yes)	*3.450****	*3.443****	*3.376****	*3.512****	*3.512****	*3.908****
	(1.121)	*(1.109)*	*(1.110)*	*(1.145)*	*(1.222)*	*(1.276)*
Control:						
District fixed effect	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*
Household Characteristics	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*
Community Characteristics	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*
Observations	1,984	1,984	1,984	1,984	1,984	1,984
R^2	0.316	0.316	0.317	0.318	0.319	0.319

The author calculated based on IFPRI- Rural Household Panel Survey Round 1

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

The income loss explains that the households living close to the rivers are more vulnerable to the adverse effects of flooding. Closer to the river, higher the chances of loss as a result of the flood, and more vulnerable they are in terms of income. The households in our sample are rural households, which in the case of Pakistan, are primarily from the agriculture sector. It means that due to such flooding events, the agriculture, livestock, and fisheries sector would be the most affected. According to the Damage Need Assessment report by the Asian Development Bank, they estimated the total losses to be approximately Pak PKR.10000 million. However, the agriculture sector alone suffered a loss of Pak PKR. 5045 million, almost half was in one section (Headhoncho, 2010).

Estimations show that there is a decrease in total losses of each household as we move away from the river. In the control group, if flood 2010 has a change of 1 percent, the log loss would have an increase of 3.440 percent

more losses. The same is the case if we see the effect in treatment groups from column 2 to 13, the percentage change in the independent variable would be the percentage change in the dependent variable. This change is positive and shows an increasing trend as we move close to the flooded area. The most loss occurred to households living at 5 km and less from the source of the flood, as 1 percent change bringing 3.908 percent more income loss and its significant at 1 percent. The total annual losses of household income, it not only just includes the damages of flood 2010 for each household. Otherwise, it is a very straight ward to test the effect on the losses due to floods. For this paper, we tried to see the total income loss to individual households with the reason that every household may or may not have the effect of loss in income due to flood as it is the main reason for losses.

Now we see the relative stability for our main effect, which we are analyzing with different control variables which we added in order to address the issue of unobserved variable bias. To answer this, we are not only relying on usually used heuristic as we used in OLS and DID estimate by not only observing the stability of our results by adding variables like the fixed effects, household characteristics, and community characteristics. We have used an additional and more recent technique of Emily Oster (Oster, 2019) for expressing for using coefficient that would be stable as a test for selection on the unobservable variables. She used this information after conducting an omitted variable bias test based on work done by Altonji, Elder, and Taber in which they assessed the effectiveness of catholic schools (Altonji et al., 2005).

Table 5a. Emily Oster estimates for distance (60 km to 35 km)

Variable	Dependent variable: Log of household losses in 2010-11						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Treatment group (=1 if household distance from flooded area)	*All*	*HH =< 60 km*	*HH =< 55 km*	*HH =< 50 km*	*HH =< 45 km*	*HH =< 40 km*	*HH =< 35 km*
Flood in 2010 (1=yes)	*3.440****	*3.447****	*3.440****	*3.432****	*3.483****	*3.525****	*3.453****
	(1.109)	*(1.110)*	*(1.109)*	*(1.114)*	*(1.107)*	*(1.118)*	*(1.122)*
Observations	*1,984*	*1,984*	*1,984*	*1,984*	*1,984*	*1,984*	*1,984*
R^2	*0.316*	*0.316*	*0.316*	*0.317*	*0.318*	*0.319*	*0.319*
Oster Estimates							
Identified set	*[3.44, 7.09]*	*[3.44, 7.09]*	*[3.45, 7.20]*	*[3.44, 7.20]*	*[3.44, 7.12]*	*[3.45, 7.21]*	*[3.44, 7.20]*
δ for $\beta = 0$ given R_{max}	*1.480*	*1.480*	*1.475*	*1.475*	*1.479*	*1.478*	*1.472*
$R_{max}(R^2 \times \Pi)$	*0.4108*	*0.4108*	*0.4121*	*0.4121*	*0.4108*	*0.4121*	*0.4108*

The author calculated based on IFPRI- Rural Household Panel Survey Round 1

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

Note:

The identified set is bounded below by $\check{\beta}$ and above by β^* calculated based on R_{max} and $\check{\delta}$ = 1

Π = 1.3

Beta calculate bound of the treatment effect

Delta calculates the relative degree of selection

Table 5b. Emily Oster estimates for distance (30 km to 05 km)

Variable	Dependent variable: Log of household losses in 2010-11					
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Treatment group (=1 if household distance from flooded area)	HH =< 30 km	HH =< 25 km	HH =< 20 km	HH =< 15 km	HH =< 10 km	HH =< 05 km
Flood in 2010 (1=yes)	*3.450****	*3.443****	*3.376****	*3.512****	*3.512****	*3.908****
	(1.121)	*(1.109)*	*(1.110)*	*(1.145)*	*(1.222)*	*(1.276)*
Observations	1,984	1,984	1,984	1,984	1,984	1,984
R^2	0.316	0.316	0.317	0.318	0.319	0.319
Oster Estimates						
Identified set	[3.43, 7.09]	[3.43, 7.09]	[3.41, 7.17]	[3.45, 7.09]	[3.61, 7.68]	[3.90, 9.83]
$\delta\ for\ \beta = 0\ given\ R_{max}$	1.48	1.48	1.45	1.49	1.51	1.39
$R_{max}(R^2\ x\ \Pi)$	0.4108	0.4108	0.4121	0.4134	0.4147	0.4147

The author calculated based on IFPRI- Rural Household Panel Survey Round 1

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

Note:

The identified set is bounded below by $\check{\beta}$ and above by β^* calculated based on R_{max} and $\check{\delta}$ = 1

$\prod$ = 1.3

Beta calculate bound of the treatment effect

Delta calculates the relative degree of selection

As shown in Table 5 (a) and 5 (b), we focus on the flood 2010 direct effect as the decline in the household income along with the change in the distance from the flooded source (distance from flood 05 km to 60 km). The test showed both the stability of the estimates. The flood treatment effect regarding the other additional key observable factors, and the results show that these observables result in tables 3, 4(a), 4(b) in explaining the income loss. These are in line with the main findings that the effect of the flood is in association with household income. As the distance of the household increases, the effect of the flood on their income would be less. On the other hand, if the households are living closer to the river or sources of flooding, the value of R^{max} is more, as the estimate of the coefficient of proportionality (suggested by Oster 2019). To summarize of our robustness of results implies that unobservable would have to be substantially more important than observables in explaining the treatment effect in order for the actual treatment effect to be zero. In the case of households living near the flood, by adding the controls even moves the estimated treatment effect further away from zero in absolute terms.

Natural disasters can have substantial and long-lasting effects on the well-being of households, especially in developing countries. Absence of formal mechanisms of insurance, tenants are likely to use multiple methods to deal with risk and its consequences. Families in developing countries face various risks like arising from weather variability, as well as from the existence of pests and epidemics, among others. They are often ill-prepared to deal with them due to the inefficiency or inexistence of formal insurance or credit mechanisms. The limited scope of public safety nets in their societies is also the other problem. For rural households, weather-related shocks are particularly threatening since the impact of hurricanes, droughts, or floods is often felt the most among the poorest in the population frequently living in rural areas. Who not only lack access to formal insurance but whose primary source of agricultural income output is severely affected by these events. As a result, households in these settings are likely to use a variety of alternative mechanisms (frequently less than optimal) to deal with natural hazards and their consequences.

4. Conclusion

This paper explored the impact of the 2010 flood on rural households' income to see the behavior as a response to an event like a flood. As we can see, the household farmers would move away from agricultural activities as an immediate response to disasters, which resulted in a decline in their income but eventually had to return to agriculture after recovering from the disaster. Repeated flooding after 2010 in 2011 and 2013 would have a further severe impact on the households which had experienced it again. For those households, farming activities

would have been more difficult as our analysis only focused on the flood event in 2010; this paper did not focus on other flooding events. Thus, while flood exposure changes the income composition of Pakistani farmers, financial assistance schemes are only short-term coping strategies. The government's need is to designs a proper mechanism, or any longer-term structural change is needed to be brought to address this issue.

In this study, we have focused on the rural household's income loss to examine the intensity of the losses of the sample households regarding the distance from the river. It will enable the government; policymakers to plan and make countermeasure strategies in terms of the construction of infrastructure to project the households living near these flooded areas in the future.

The empirical results carry significant implications, especially for the developing countries which experience frequent natural disasters. We need to look forward to the agents who can contribute to protecting these vulnerable households during disasters, especially the ones living near to the flood sources like 05 km to 20 km as the impact on them is more than the once-living far. The effects these events become more for low-income countries such as Pakistan. As the prime activity for income for these countries is agriculture so the farmers should be the main focus when planning development projects after disasters so to meet an urgent need for there subsistence, and if not, they may not be able to cope up in future with these disasters.

The effect of natural disasters in Pakistan have an impact which is long-lasting on the income of the affected households, but more on the ones living close to the rivers. Especially living from 5 km to 20 km for these households, the coping strategies need to be different. While the farmers are more vulnerable to income vulnerably due to seasonal climatic risks as the frequency and magnitude of natural disasters change. For this reason, we need to emphasize on strengthening, developing, and adapting to the capability of farmers to combat natural disasters. The empirical analysis suggests that households after flooding tend to move away from activities related to agriculture due to which there is a decline in their income. Therefore, the observed changes in income strategies do not necessarily imply a structural change; instead, they reflect flood-affected household's short-term coping with the harms of disaster.

We provide a regional perspective, and we suggest to both donors and Sahelian countries a regional approach to the problem. Since regional climatic shocks are often of great magnitude and affecting multiple countries at the same time, the response goes beyond the capacity of individual countries and calls for a regional approach and stronger coordination. Our results suggest that rural households that are being affected by natural disasters their income are more vulnerable to these shocks, especially floods in this case. The effect of disaster is causing more income losses to these households, which are agricultural households as the effect of floods are far more than of another sector of the economy. The overall income of people involved directly or indirectly is also affected due to such shocks and spill away effect. Other aspects like commitments of government need to be analyzed as well while focusing on a human resource like, for example, employees working for the government sector; their commitment should be an essential variable as well to eradicating household vulnerability for rural areas in Pakistan (Ajmal, 2019).

The findings in this paper can give insight for government and donor agencies to see important implications for planning targeted interventions such as social safety nets, as well as other forms of short- or long-term support for vulnerable households. Firstly, for many countries, especially developing countries, which are trying to relieve programs based on estimates which they obtained from populations at other points of time. Therefore, these estimates will be more likely to be focused on that events point of time when that survey or estimation was conducted. The method we proposed and used for estimations provides a more complete, fast, and reliable perspective to the extend for relief program planning. Secondly, a better-targeted program to address vulnerable households by distinguishing among the one based on their needs would be more productive and substantial.

Acknowledgments

References

Ajmal, A. (2019, July 6). Measuring the Level of Commitment in Tertiary Child Health Care Units for Effective Performance in Pakistan. *The European Conference on Psychology & the Behavioral Sciences 2019: Official Conference Proceedings*. The European Conference on Psychology & the Behavioral Sciences 2019, The Jurys Inn Brighton Waterfront, Brighton, United Kingdom. Retrieved from https://papers.iafor.org/submission51273/

Alfani, F., Dabalen, A., Fisker, P., & Molini, V. (2019). Vulnerability to stunting in the West African Sahel.

Food Policy, *83*, 39–47. https://doi.org/10.1016/j.foodpol.2018.11.002

Anttila-Hughes, J., & Hsiang, S. (2013). *Destruction, Disinvestment, and Death: Economic and Human Losses Following Environmental Disaster* (SSRN Scholarly Paper ID 2220501). Social Science Research Network. Retrieved from https://papers.ssrn.com/abstract=2220501

Azeem, M. M., Mugera, A. W., & Schilizzi, S. (2018). Vulnerability to Multi-Dimensional Poverty: An Empirical Comparison of Alternative Measurement Approaches. *The Journal of Development Studies*, *54*(9), 1612–1636. https://doi.org/10.1080/00220388.2017.1344646

Barnett, B. J., & Mahul, O. (2007). Weather Index Insurance for Agriculture and Rural Areas in Lower-Income Countries. *American Journal of Agricultural Economics*, *89*(5), 1241–1247. https://doi.org/10.1111/j.1467-8276.2007.01091.x

Baulch, B., & Hoddinott, J. (2000). Economic mobility and poverty dynamics in developing countries. *The Journal of Development Studies*, *36*(6), 1–24. https://doi.org/10.1080/00220380008422652

Briguglio, L., Cordina, G., Farrugia, N., & Vella, S. (2008). *Profiling economic vulnerability and resilience in small states: Conceptual underpinnings*. Retrieved from https://www.um.edu.mt/library/oar/handle/123456789/18562

Calvo, C. (2018). Vulnerability to poverty: Theoretical approaches. *Handbook of Research on Economic and Social Well-Being*. Retrieved from https://www.elgaronline.com/view/edcoll/9781781953709/9781781953709.00016.xml

Celidoni, M. (2013). Vulnerability to poverty: An empirical comparison of alternative measures. *Applied Economics*, *45*(12), 1493–1506. https://doi.org/10.1080/00036846.2011.624271

Chambers, R. (1989). *Editorial Introduction: Vulnerability, Coping and Policy*. https://doi.org/10.1111/j.1759-5436.1989.mp20002001.x

Chaudhuri, S. (2003). Assessing vulnerability to poverty: Concepts, empirical methods and illustrative examples. *Department of Economics, Columbia University, New York*, 56.

Damage Need Assessment (DNA) Report of Asian Development Bank, Nov. (2010). Data Archive of Flood Forecasting Division, Pakistan.

Dercon, S. (2002). Income Risk, Coping Strategies, and Safety Nets. *The World Bank Research Observer*, *17*(2), 141–166. https://doi.org/10.1093/wbro/17.2.141

Dercon, S. (2006). Economic reform, growth and the poor: Evidence from rural Ethiopia. *Journal of Development Economics*, *81*(1), 1–24. https://doi.org/10.1016/j.jdeveco.2005.05.008

Dutta, I., Foster, J., & Mishra, A. (2011). On measuring vulnerability to poverty. *Social Choice and Welfare*, *37*(4), 743. https://doi.org/10.1007/s00355-011-0570-1

EM-DAT (2019). (n.d.). Retrieved November 13, 2019, from www.emdat.be, Brussels, Belgium

Epping-Jordan, J. E., van Ommeren, M., Ashour, H. N., Maramis, A., Marini, A., Mohanraj, A., Noori, A., Rizwan, H., Saeed, K., Silove, D., Suveendran, T., Urbina, L., Ventevogel, P., & Saxena, S. (2015). Beyond the crisis: Building back better mental health care in 10 emergency-affected areas using a longer-term perspective. *International Journal of Mental Health Systems*, *9*(1), 15. https://doi.org/10.1186/s13033-015-0007-9

Feeny, S., & McDonald, L. (2016). Vulnerability to Multidimensional Poverty: Findings from Households in Melanesia. *The Journal of Development Studies*, *52*(3), 447–464. https://doi.org/10.1080/00220388.2015.1075974

FFC. (2011). Federal Flood Commission of Pakistan, Annual Flood Report-2010.

Food and Agriculture Organization (FAO). (2015). *The impact of disasters on agriculture and food security*. United Nations.

Glewwe, P., & Hall, G. (1998). Are some groups more vulnerable to macroeconomic shocks than others? Hypothesis tests based on panel data from Peru. *Journal of Development Economics*, *56*(1), 181–206. https://doi.org/10.1016/S0304-3878(98)00058-3

Hansen, J., Hellin, J., Rosenstock, T., Fisher, E., Cairns, J., Stirling, C., Lamanna, C., van Etten, J., Rose, A., & Campbell, B. (2019). Climate risk management and rural poverty reduction. *Agricultural Systems*, *172*, 28–46. https://doi.org/10.1016/j.agsy.2018.01.019

Heeks, R., & Ospina, A. V. (2019). Conceptualising the link between information systems and resilience: A developing country field study. *Information Systems Journal*, *29*(1), 70–96. https://doi.org/10.1111/isj.12177

Heeringa, S. G., West, B. T., Berglund, P. A., West, B. T., & Berglund, P. A. (2017). Applied Survey Data Analysis. Chapman and Hall/CRC. https://doi.org/10.1201/9781315153278

Hoddinott, J., & Quisumbing, A. (2010). Methods for Microeconometric Risk and Vulnerability Assessment. In R. Fuentes-Nieva, & P. A. Seck (Eds.), *Risk, Shocks, and Human Development: On the Brink* (pp. 62–100). Palgrave Macmillan UK. https://doi.org/10.1057/9780230274129_4

Hoddinott, J., & Quisumbing, A. (n.d.). Social Protection Discussion Paper Series. *2003*, 78.

JAXA Global Rainfall Watch (GSMaP). (n.d.). Retrieved May 29, 2019, from https://sharaku.eorc.jaxa.jp/GSMaP/

Kamanou, G., & Morduch, J. (2002). *Measuring vulnerability to poverty* (Working Paper 2002/58). WIDER Discussion Paper. Retrieved from https://www.econstor.eu/handle/10419/53096

Kreft, S., Eckstein, D., Dorsch, L., & Fischer, L. (2016). Global Climate Risk Index 2016: Who Suffers Most from Extreme Weather Events? Weather-related Loss Events in 2014 and 1995 to 2014 Germanwatch eV, Bonn, Germany.

Kurosaki, T., & Fafchamps, M. (2002). Insurance market efficiency and crop choices in Pakistan. *Journal of Development Economics*, *67*(2), 419–453. https://doi.org/10.1016/S0304-3878(01)00188-2

Kutner, M. H. (Ed.). (2005). Applied linear statistical models (5th ed.). McGraw-Hill Irwin.

Lechner, M., Rodriguez-Planas, N., & Fernández Kranz, D. (2016). Difference-in-difference estimation by FE and OLS when there is panel non-response. *Journal of Applied Statistics*, *43*(11), 2044–2052. https://doi.org/10.1080/02664763.2015.1126240

Ligon, E., & Schechter, L. (2003). Measuring Vulnerability. *The Economic Journal*, *113*(486), C95–C102. https://doi.org/10.1111/1468-0297.00117

Maccini, S., & Yang, D. (2009). Under the Weather: Health, Schooling, and Economic Consequences of Early-Life Rainfall. *American Economic Review*, *99*(3), 1006–1026. https://doi.org/10.1257/aer.99.3.1006

Nazli, H., & Haider, S. H. (n.d.). *Pakistan Rural Household Panel Survey 2012 (Round 1): Methodology and Community Characteristics*. 45.

Neter et al. (1996). *Applied linear statistical models* (4th ed.). Irwin, Chicago.

Oster, E. (2019). Unobservable Selection and Coefficient Stability: Theory and Evidence. *Journal of Business & Economic Statistics*, *37*(2), 187–204. https://doi.org/10.1080/07350015.2016.1227711

Pritchett, L. (2000). *Quantifying Vulnerability to Poverty: A Proposed Measure, with Application to Indonesia*. 32.

Qiu, L.-Y., & He, L.-Y. (2017). Can Green Traffic Policies Affect Air Quality? Evidence from A Difference-in-Difference Estimation in China. *Sustainability*, *9*(6), 1067. https://doi.org/10.3390/su9061067

Rossouw, S. (2017). Measuring the vulnerability of sub-national regions: Integrating relative location. *South African Journal of Economic and Management Sciences*, *20*(1). https://doi.org/10.4102/sajems.v20i1.1766

Singleton, G. R., & Gregory, R. (2007). *Geologic Storage of carbon dioxide: Risk analyses and implications for public acceptance* [Thesis, Massachusetts Institute of Technology]. Retrieved from https://dspace.mit.edu/handle/1721.1/40378

Snow, J. (1855). *On the Mode of Communication of Cholera*. John Churchill.

Vieider, F. M., Martinsson, P., Nam, P. K., & Truong, N. (2019). Risk preferences and development revisited. *Theory and Decision*, *86*(1), 1–21. https://doi.org/10.1007/s11238-018-9674-8

Wallquist, L., Visschers, V. H. M., & Siegrist, M. (2010). Impact of Knowledge and Misconceptions on Benefit and Risk Perception of CCS. *Environmental Science & Technology,* *44*(17), 6557–6562. https://doi.org/10.1021/es1005412

Ward, P. S. (2016). Transient Poverty, Poverty Dynamics, and Vulnerability to Poverty: An Empirical Analysis Using a Balanced Panel from Rural China. *World Development*, *78*, 541–553. https://doi.org/10.1016/j.worlddev.2015.10.022

Wisner, B., Blaikie, P. M., Blaikie, P., Cannon, T., & Davis, I. (2004). *At risk: Natural hazards, people's vulnerability and disasters* (2nd ed.). London Routledge.

Zhang, Y., & Wan, G. (2009). How Precisely Can We Estimate Vulnerability to Poverty? *Oxford Development Studies*, *37*(3), 277–287. https://doi.org/10.1080/13600810903094471

Notes

Data is available on http://www.emdat.be/natural-disasters-trends (accessed 30 April 2019). In interpreting such data, we should pay attention to the possibility that the reported increase is partly due to an increased tendency to report, not necessarily an increase in the occurrence of disasters.
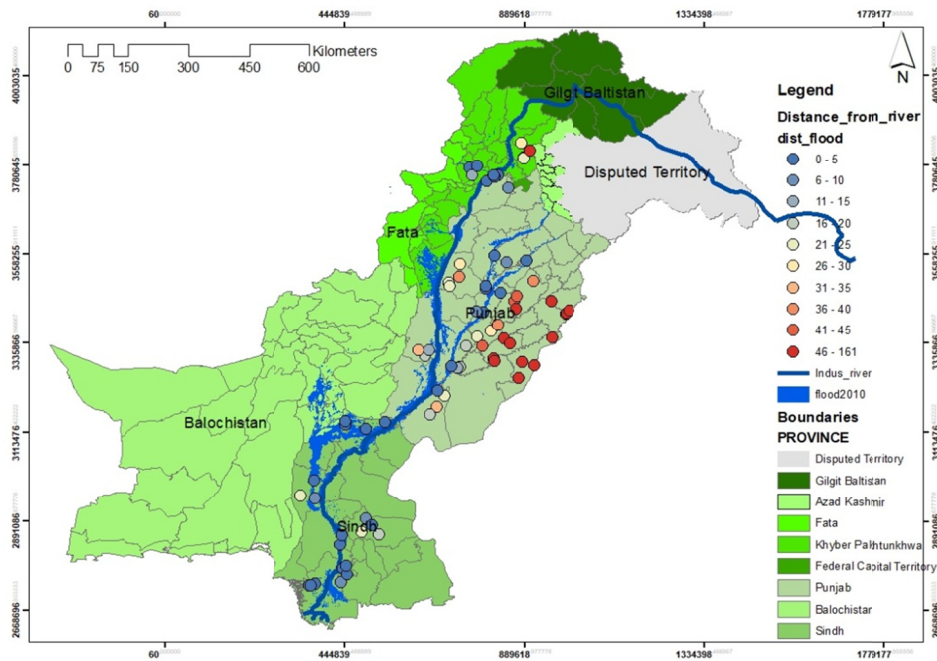
Appendix A

Using natural shapefile in ArcGIS contains information about the river, waterways, forests, parks, and other parts of Pakistan. We obtained data files from the open-source of Geofabrik it provides OpenStreetMap services and based in Karlsruhe, Germany. Appendix figure 2 shows the map location and flooded areas of Pakistan. The flood information used to calculate the distance of each household from the nearest river.
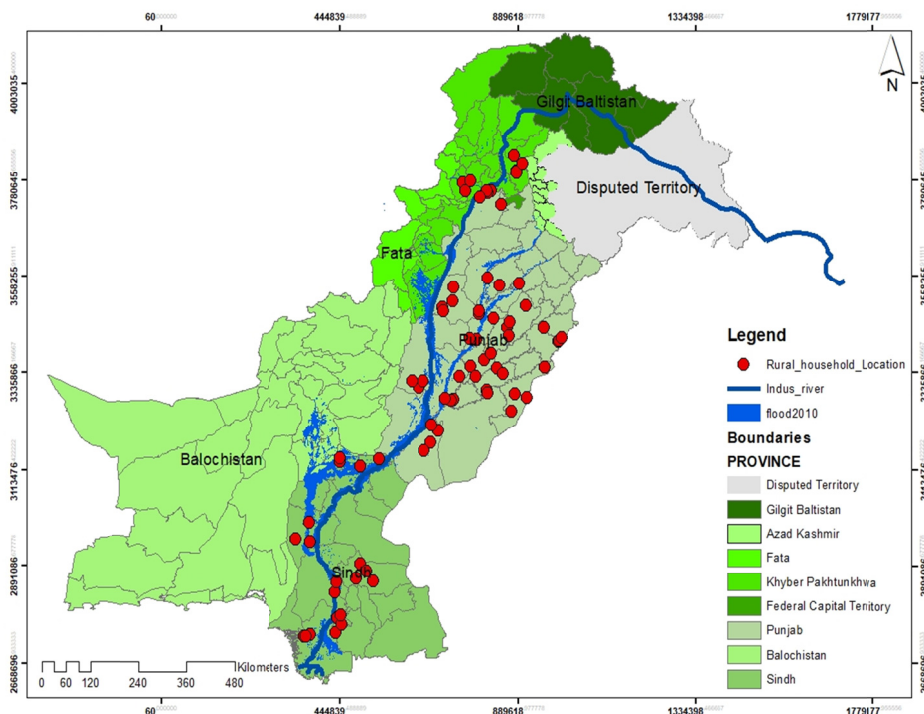
Appendix Figure 1. Location of household from RHPS

Source: RHPS and Pakistan Administrative data from Geofabrik

The distance calculated by using near tool from proximity from Arc Toolbox. The distance of each household calculated in the attribute table, which was exported to an excel file to use a variable for calculating the distance.

Appendix figure 2. Village location from the distance from the river

Source: RHPS and Pakistan Administrative data from Geofabrik
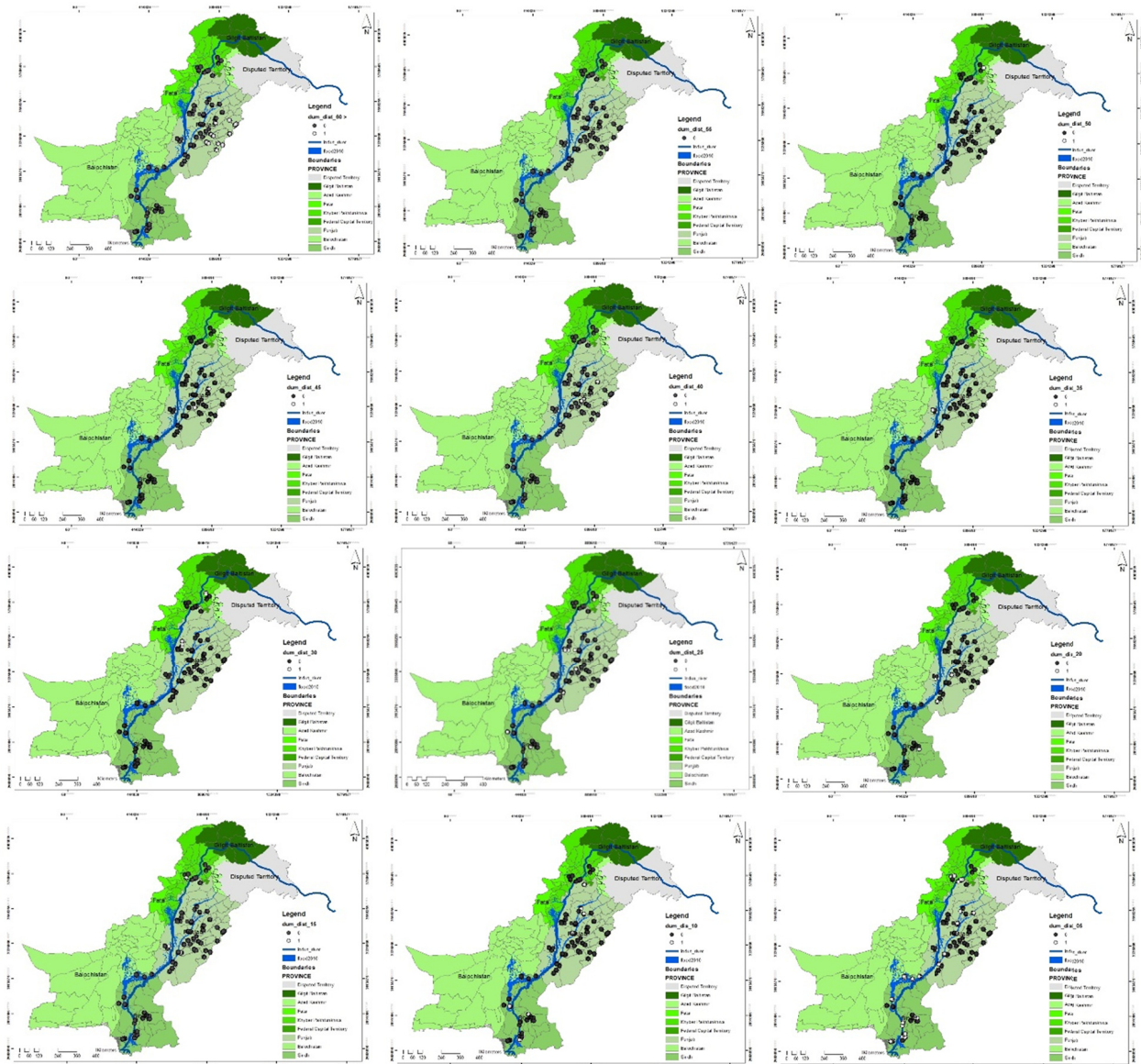
Appendix figure 3. Calculated the distance from the flood

Source: RHPS and Pakistan Administrative data from Geofabrik

Copyrights

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.