

Assessing the flood exposure and structural vulnerability in explaining regional welfare differences in Pakistan

Wareesha Malik

wareeshamalik.ds@gmail.com

COMSATS University Islamabad

Article

Keywords: Flood exposure, Structural vulnerability, Regional welfare, Consumption, Shapley decomposition, Pakistan

Posted Date: September 18th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-10946384/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

Assessing the flood exposure and structural vulnerability in explaining regional welfare differences in Pakistan

Wareesha Malik*

Independent Researcher, Abbottabad, Pakistan

*Corresponding Author: Wareesha Malik (Email: wareeshamalik.ds@gmail.com)

Abstract

Global estimates indicate that low- and middle-income countries account for nearly 89% of the world's flood-exposed population, yet the relative importance of flood exposure and structural vulnerability in explaining regional welfare differences remains poorly understood. This study examined the associations and relative importance of population flood exposure and multidimensional structural vulnerability across 127 districts of Pakistan. District-level data were analysed using ordinary least squares regression with province fixed effects and Conley spatial heteroskedasticity and autocorrelation consistent (HAC) standard errors. Shapley value decomposition was used to quantify the relative importance of population flood exposure, six structural vulnerability domains, and long-term changes in temperature and rainfall in explaining two welfare outcomes: average district consumption and average consumption among poor households. Structural vulnerability was negatively associated with both welfare measures. For average district consumption, the six structural vulnerability domains collectively accounted for more than 90% of the explained variation, whereas population flood exposure contributed less than 1%. For consumption among poor households, long-term temperature change was the largest individual contributor, although the structural vulnerability domains collectively explained a substantially larger share of the variation. Overall, population flood exposure alone provided limited insight into regional welfare differences, while structural vulnerability emerged as the principal explanation. These findings suggest that integrating structural vulnerability alongside flood exposure may strengthen climate adaptation and disaster risk reduction planning.

Keywords: Flood exposure; Structural vulnerability; Regional welfare; Consumption; Shapley decomposition; Pakistan

1. Introduction

Climate change is increasing global temperatures and altering precipitation patterns, contributing to more frequent and intense climate-related hazards worldwide [1–3]. Among these hazards, floods affect more people than any other natural hazard and are associated with substantial economic and welfare losses [4,5]. Recent estimates indicate that approximately 1.81 billion people, representing nearly 23% of the global population, live in areas exposed to 1-in-100-year floods, of whom 1.24 billion reside in South and East Asia. A 100-year flood corresponds to a 1% probability of occurring in any given year, a 10% probability over a decade, and approximately a 50% probability over a 68-year period. Moreover, low- and middle-income countries account for 89% of the world's flood-exposed population [6]. Population flood exposure is therefore increasingly used in international disaster-risk and climate-adaptation assessments to identify populations potentially exposed to flooding and inform resource prioritisation. Yet population exposure alone provides limited information about why welfare differs across places.

The welfare implications of population flood exposure can differ because populations are embedded in different structural contexts. Housing, livelihoods, education, health, infrastructure, and demographic

conditions shape the capacity of populations to anticipate, withstand, and respond to environmental risks [7–12]. A substantial body of research has documented welfare losses associated with environmental and climate risks, including declines in household consumption, income losses, reduced agricultural productivity, and increased poverty, particularly among populations with limited capacity to cope with environmental stress [13–17]. Households with limited assets, inadequate access to insurance and credit, and weak social protection may face greater difficulty recovering from environmental shocks, increasing the potential for persistent welfare losses and poverty traps [11,12,17]. Cross-country evidence further suggests that stronger structural conditions can reduce welfare losses associated with environmental risks even where population exposure remains high [18,19].

Despite this conceptual distinction, relatively little empirical research has directly compared the relative importance of population flood exposure and multidimensional structural vulnerability in explaining spatial differences in welfare. Existing flood-risk studies have predominantly characterised exposure through the physical distribution and intensity of flood hazards, while comparatively less attention has been given to the population dimension of exposure. Rentschler et al. [6] demonstrated at the global scale that combining population flood exposure with poverty information reveals patterns that exposure counts alone do not capture. Their analysis highlighted the importance of considering the characteristics of exposed populations alongside exposure itself. Building on this perspective, the present study examines population flood exposure alongside multidimensional structural vulnerability at the subnational scale, allowing their relative explanatory importance for welfare differences to be assessed within a common framework.

Pakistan provides an informative setting for this analysis because it combines substantial population flood exposure with pronounced spatial differences in structural vulnerability and welfare. Approximately 26.9% of Pakistan's population falls within the modelled 100-year flood inundation footprint, indicating that a substantial share of the population is potentially exposed under this severe flood scenario. This national estimate, however, masks considerable spatial variation, with population exposure concentrated largely along the Indus River system and across parts of Punjab and Sindh [6]. Pakistan also faces substantial climate-related pressures, including rising temperatures and changing precipitation patterns that affect climate-sensitive livelihoods and economic systems [1,20,21]. Its geography, including the Indus River Basin and dependence on monsoon rainfall, glacier and snowmelt, and complex river systems, creates substantial potential for flood exposure [22–25].

Pakistan's historical experience also illustrates the broader significance of flood risk within the country. According to the Government of Pakistan's *Fiscal Risk Statement*, the country experienced 224 major natural disasters between 1980 and 2024, including 109 floods that affected more than 100 million people and caused an estimated US\$36.4 billion in damages [25]. These historical impacts provide important context for the country's flood-risk landscape, although the present study does not examine the welfare consequences of individual flood events. Instead, it focuses on a long-term measure of population flood exposure, defined as the proportion of the district population located within a modelled 100-year flood inundation zone. This measure captures persistent spatial differences in potential population exposure and provides a nationally consistent basis for comparing districts. At the same time, substantial differences in housing conditions, livelihood opportunities, education, health services, infrastructure, and demographic characteristics contribute to considerable variation in welfare and structural vulnerability across Pakistan [20].

Previous studies in Pakistan have documented associations between floods, environmental change, and household welfare [26–28], while other research has assessed multidimensional vulnerability in flood-prone areas and highlighted the importance of housing, livelihoods, access to basic services, and demographic characteristics [20,29,30]. Much of this evidence, however, focuses on individual flood events, specific provinces or communities, or particular dimensions of vulnerability. Comparatively few studies jointly

examine population flood exposure and multidimensional structural vulnerability across Pakistan's districts, and fewer still assess their relative explanatory importance for welfare.

This study addresses this gap by examining the associations between population flood exposure, multidimensional structural vulnerability, and welfare across 127 districts of Pakistan. Welfare is measured using two complementary consumption-based outcomes: mean monthly per adult-equivalent consumption and mean monthly per adult-equivalent consumption among poor households. Structural vulnerability is measured using the District Vulnerability Index for Pakistan (DVIP), comprising six non-monetary domains: housing, communication and transport, livelihoods, health, education, and demographics [32]. The analysis estimates associations while accounting for province-level differences and spatial dependence and controlling for long-term changes in temperature and rainfall. Shapley value decomposition is subsequently used to quantify the relative importance of population flood exposure, multidimensional structural vulnerability, and climatic variables in explaining variation in each welfare outcome [37,38].

By bringing these dimensions together within a common subnational framework, this study makes four contributions. First, it examines population flood exposure as a human-centred measure of environmental exposure and considers it alongside the structural characteristics of the places in which exposed populations live. Second, it moves beyond income-based poverty measures by examining a broader set of non-monetary structural conditions, including housing, livelihoods, education, health, communication and transport, and demographics. Third, it compares two complementary consumption-based welfare outcomes—overall district consumption and consumption among poor households—allowing the spatial correlates of welfare to be examined across different segments of the welfare distribution. Finally, it quantifies the relative explanatory importance of population flood exposure, multidimensional structural vulnerability, and climatic factors within a nationally consistent district-level framework. Together, these contributions provide a spatially differentiated perspective on how population exposure and persistent structural conditions relate to welfare differences and offer evidence relevant to climate adaptation and development planning in hazard-prone settings.

2. Methodology

2.1 Study Area and Unit of Analysis

Pakistan provides an informative setting for this study because it exhibits substantial subnational differences in population flood exposure, structural conditions, and regional welfare outcomes [1,31]. Pakistan comprises four main provinces—Punjab, Sindh, Khyber Pakhtunkhwa, and Balochistan—which differ considerably in climatic conditions, livelihoods, infrastructure, and socioeconomic characteristics, resulting in pronounced variation in both population flood exposure and multidimensional structural vulnerability. The unit of analysis is the administrative district. This study covers the 145 districts located within these four provinces. Of these, 18 districts were excluded because of missing climate or consumption data, yielding a final analytical sample of 127 districts: 36 in Punjab, 29 in Sindh, 35 in Khyber Pakhtunkhwa, and 27 in Balochistan. Districts were selected because they represent the finest spatial scale at which all datasets could be consistently integrated and constitute an important administrative level for planning and implementing development, disaster risk reduction, and social protection programmes. They also capture important spatial heterogeneity in population flood exposure, multidimensional structural vulnerability, and regional welfare that would be masked at the provincial level [32].

2.2 Data Sources

Regional welfare. District-level consumption data were obtained from the Pakistan Poverty Alleviation Mapping exercise undertaken by the World Bank and the Pakistan Poverty Alleviation Fund (PPAF) and

can found at: <https://datacatalog.worldbank.org/search/dataset/0066822/pakistan-poverty-map-2019-2020>. The estimates are based on the Pakistan Social and Living Standards Measurement Survey (PSLM) 2019–20 and were generated using small-area estimation techniques. Welfare was measured using average consumption expenditure, which is widely used as a proxy for welfare because it is generally less sensitive than income to seasonal fluctuations and reporting bias [16,33,34,35]. Two outcome variables were analysed: the natural logarithm of mean monthly per adult equivalent consumption and the natural logarithm of mean monthly per adult equivalent consumption among poor.

Multidimensional structural vulnerability. Structural vulnerability was measured using the District Vulnerability Index for Pakistan (DVIP), developed by the Population Council Pakistan and endorsed by the Government of Pakistan [32]. Available at: <https://dvip.pcpak.org/DVIP-Report-2025.pdf>. The DVIP comprises 21 indicators grouped into six non-monetary domains: housing, communication and transport, livelihoods, access to health services, access to public education, and demographics (Fig. 1). Because the DVIP is constructed exclusively from non-monetary indicators, it does not include income, consumption, or poverty measures, thereby reducing conceptual overlap with the study's consumption measures. The index was constructed using multiple nationally representative datasets, including the Pakistan Mouza Census 2020, Pakistan Social and Living Standards Measurement Survey (PSLM) 2019–20, Pakistan Institute for Education (PIE) School Directory 2020–21, and the Population and Housing Census 2023. Indicators were aligned so that higher values represented greater vulnerability, transformed to district percentile scores, and aggregated using equal weights across indicators and domains. The composite index demonstrated high internal consistency (Cronbach's $\alpha = 0.92$). Both the composite index and the six domain-specific scores were obtained directly from the published DVIP dataset. Using an externally developed and nationally standardised index also avoids researcher-driven decisions regarding indicator selection and weighting, thereby enhancing the transparency, reproducibility, and policy relevance of the analysis. The DVIP follows the equal-weighting scheme adopted in its original construction. Although alternative weighting approaches could produce different composite scores, the domain-level Shapley decomposition provides an independent assessment of the relative contribution of each vulnerability domain to regional welfare.

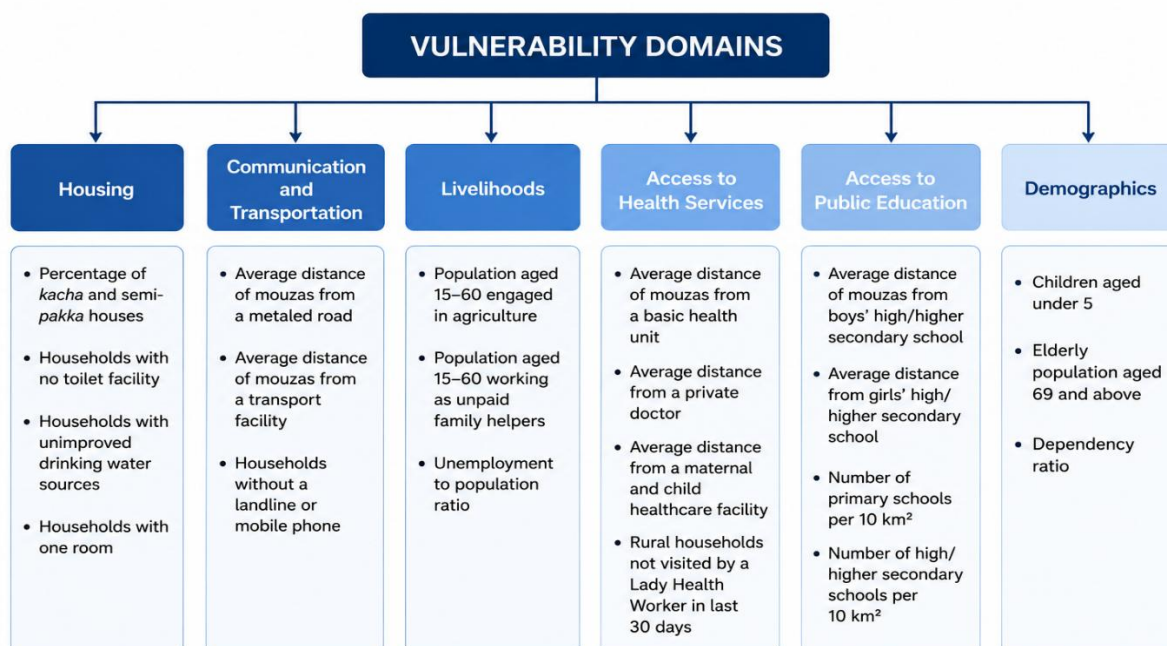


Fig. 1. Domains of Vulnerability and their Indicators

Flood exposure. District-level flood exposure was measured as the percentage of the district population exposed to a modelled 100-year return-period flood inundation zone with a minimum water depth of 30 cm. This is a population-based measure that quantifies the share of district residents potentially exposed to the modelled flood footprint, rather than the geographic area of the district subject to inundation. Measuring population exposure rather than inundated area is important because the potential human consequences of flooding depend not only on the extent of inundation but also on where people are located. A large inundated area with sparse population may therefore imply a different potential welfare burden than a smaller inundated area containing a large share of the population. By quantifying the proportion of residents within the modelled flood footprint, the measure provides a more directly welfare-relevant basis for comparing potential population exposure across districts than area-based measures alone. The indicator was derived from the Humanitarian Data Exchange (HDX) using the GAIA risk assessment pipeline developed by the Heidelberg Institute for Geoinformation Technology (HeiGIT), which combines Joint Research Centre (JRC) Global Flood Maps with WorldPop population data. Available at: <https://data.humdata.org/dataset/pakistan---risk-assessment-indicators>. A standardised 100-year flood hazard footprint was used to represent long-term geographic flood exposure rather than the impacts of individual flood events.

Long-term climate change. Long-term climatic change was represented by district-level changes in mean annual temperature and rainfall derived from the Climate Research Unit (CRU TS4) observational climate dataset. Temperature and rainfall change were calculated as the difference between average annual conditions during 1970–1980 and 2006–2016. These variables represent long-term climatic trends rather than short-term climate variability and were obtained from the DVIP dataset, which processed the CRU TS4 observations to district level. Available at: <https://dvip.pcpak.org/DVIP-Report-2025.pdf>.

Province fixed effects. Province fixed effects were included in all regression models using binary indicators for Sindh, Khyber Pakhtunkhwa, and Balochistan, with Punjab specified as the reference

category. These variables account for broad province-level differences that may be associated with district-level welfare outcomes.

2.3 Data Analysis

Table 1 summarises the variables included in the statistical analysis. Descriptive statistics were calculated to summarise the distribution of all variables across the 127 districts. Choropleth maps were produced to visualise the spatial distribution of flood exposure, structural vulnerability, and regional welfare. Scatter plots with fitted regression lines were used to examine the relationships between the welfare outcomes and the principal explanatory variables and to assess the assumption of linearity. Pearson correlation coefficients and variance inflation factors (VIFs) were examined to assess multicollinearity among the explanatory variables before regression modelling. The associations between flood exposure, structural vulnerability, and regional welfare were estimated using ordinary least squares (OLS) regression with province fixed effects and Conley spatial heteroskedasticity- and autocorrelation-consistent (HAC) standard errors. Shapley value decomposition was subsequently applied to quantify the relative contribution of the explanatory variables to the explained variation in each welfare outcome.

Table 1 Variable Definitions, Measurement, and Role in Analysis

Variable	Definition / Measurement	Measurement Type	Role in Model
Consumption	Log transformed mean monthly per adult equivalent consumption per district (Rs.)	Continuous	Dependent - Model A
Consumption among the Poor	Log transformed mean monthly per adult equivalent consumption of poor households per district (Rs.)	Continuous	Dependent - Model B
Flood Exposure (%)	Percentage of district population exposed to 100-year return period flooding	Continuous	Key predictor
Vulnerability Index (Composite)	Composite DVIP index (0-1 scale); equal weighted aggregation of six domains	Continuous	Key predictor
Temperature Change (°C)	District level change in mean annual temperature between 1970-80 and 2006-16	Continuous	Control
Rainfall Change (mm)	District level change in mean annual precipitation between 1970-80 and 2006-16	Continuous	Control
Province Dummies	Binary indicators for Sindh, Khyber Pakhtunkhwa, and Balochistan (Punjab = reference)	Binary/Categorical	Control

2.4 OLS Regression with Spatially Robust Standard Errors

Ordinary least squares (OLS) regression with province fixed effects was used to estimate the associations between population flood exposure, structural vulnerability, and regional welfare. Because the objective of this study was to examine district-level associations rather than identify causal effects, the estimated coefficients are interpreted as conditional associations. For each welfare outcome, two regression models were estimated. The baseline models (A1 and B1) included population flood exposure, long-term temperature change, long-term rainfall change, and province fixed effects. The full models (A2 and B2) additionally included the composite District Vulnerability Index for Pakistan (DVIP) to assess whether multidimensional structural vulnerability explained regional welfare beyond population flood exposure and long-term climatic conditions.

Prior to model estimation, scatter plots with fitted regression lines were examined to assess linearity, while Pearson correlation coefficients and variance inflation factors (VIFs) were used to evaluate multicollinearity. The dependent variables were natural log-transformed to reduce positive skewness. When the six structural vulnerability domains were entered simultaneously, VIF values exceeded five, indicating moderate multicollinearity and limiting the interpretation of individual regression coefficients. Consequently, the primary regression models used the composite DVIP score, whereas the individual vulnerability domains were evaluated separately using Shapley value decomposition.

Because geographically proximate districts may exhibit correlated regression errors, statistical inference was based on Conley (1999) spatial heteroskedasticity- and autocorrelation-consistent (HAC) standard errors [36]. Rather than adopt a conventional but arbitrary cutoff, the spatial bandwidth was selected using a data-driven diagnostic based on the empirical covariogram of the OLS residuals, following Lehner (2026), which identifies the approximate spatial dependence range from the distance at which residual covariance approaches or crosses zero. The estimated zero-crossing distances were approximately 160 km (A1), 115 km (A2), 120 km (B1), and 110 km (B2). Based on these diagnostics, a common bandwidth of 175 km was applied across all four specifications to provide a consistent and conservative adjustment for spatial dependence. The Conley estimator used an Epanechnikov kernel and Haversine great-circle distances.

The models are specified as:

Model A1:

$$Y_i = \alpha + \beta_1 F_i + \beta_2 T_i + \beta_3 R_i + \sum_{k=1}^3 \gamma_k P_{k,i} + \varepsilon_i$$

Model B1:

$$Z_i = \alpha + \beta_1 F_i + \beta_2 T_i + \beta_3 R_i + \sum_{k=1}^3 \gamma_k P_{k,i} + \varepsilon_i$$

Model A2:

$$Y_i = \alpha + \beta_1 F_i + \beta_2 T_i + \beta_3 R_i + \beta_4 V_i + \sum_{k=1}^3 \gamma_k P_{k,i} + \varepsilon_i$$

Model B2:

$$Z_i = \alpha + \beta_1 F_i + \beta_2 T_i + \beta_3 R_i + \beta_4 V_i + \sum_{k=1}^3 \gamma_k P_{k,i} + \varepsilon_i$$

where subscript i denotes district; Y_i is the natural log of mean consumption per adult equivalent; Z_i is the natural log of mean consumption among poor; F_i is population flood exposure (%); T_i is temperature change ($^{\circ}\text{C}$); R_i is rainfall change (mm); V_i is the composite DVIP index; $P_{k,i}$ are province fixed effects (binary indicators for Sindh, Khyber Pakhtunkhwa, and Balochistan, with Punjab as the reference); α is the intercept; β_1 through β_4 are regression coefficients; γ_k are province fixed effect coefficients; and ε_i is the error term.

2.5 Relative Importance Regression

To quantify the relative contribution of population flood exposure, long-term climatic change, and multidimensional structural vulnerability to regional welfare, Shapley value decomposition was applied using the lmg metric. Shapley decomposition partitions the model's explained variance (R^2) by calculating the average marginal contribution of each predictor across all possible model orderings, thereby providing an order-independent measure of relative importance [37,38]. Unlike standardized regression coefficients, Shapley decomposition accounts for shared explanatory power among correlated predictors, making it particularly suitable when explanatory variables exhibit moderate intercorrelations [38].

Separate decompositions were performed for the two welfare outcomes. The analysis included population flood exposure, long-term temperature change, long-term rainfall change, and the six structural vulnerability domains. Province fixed effects were not included in the decomposition because the objective was to compare the explanatory contribution of the substantive predictors rather than to partition variation attributable to broad administrative differences. The resulting Shapley values represent the percentage contribution of each predictor to the model's explained variance (R^2) and should be interpreted as measures of relative explanatory importance rather than causal effects.

3. Results

3.1 Descriptive Statistics

Table 2 summarises the distribution of the study variables across the 127 districts included in the analysis. Mean monthly per adult-equivalent consumption ranged from Rs. 3,282 to Rs. 11,316, indicating substantial variation in regional welfare. Mean consumption among poor exhibited a narrower range than average consumption. Population flood exposure varied markedly across districts, ranging from 0% to 96.4% of the population. Long-term temperature change averaged 0.61°C , whereas rainfall change ranged from -57.7 mm to 256.7 mm. The six structural vulnerability domains showed similar distributions, with mean values close to 0.49 and standard deviations of approximately 0.29, reflecting the standardized construction of the District Vulnerability Index for Pakistan (DVIP).

Table 2 Descriptive Statistics of Key Variables (N = 127 districts)

Variable	Mean	SD	Min	Max
Mean Per Adult-Equivalent Consumption (Rs/month)	5470.84	1318.54	3281.88	11316.04
Mean Per Adult-Equivalent Consumption of the poor (Rs/month)	3162.58	117.63	2837.95	3381.35
Flood exposure (% , 100-year)	27.54	31.16	0.00	96.39

Change in temperature (°C)	0.61	0.32	0.22	1.38
Change in rainfall (mm)	41.37	59.19	-57.70	256.70
Structural vulnerability (DVIP domains)				
Housing	0.494	0.292	0.000	0.992
Communication & Transport	0.489	0.285	0.000	1.000
Livelihoods	0.491	0.291	0.000	1.000
Health	0.489	0.287	0.000	1.000
Education	0.490	0.286	0.000	0.984
Demographics	0.493	0.291	0.000	1.000
Province distribution (n districts)				
Punjab (reference)	36	—	—	—
Sindh	29	—	—	—
Khyber Pakhtunkhwa (KP)	35	—	—	—
Balochistan	27	—	—	—

Note: SD = standard deviation. Structural vulnerability domains are standardized scores ranging from 0 to 1.

3.2 Spatial Distribution of Flood Exposure, Vulnerability and Poverty

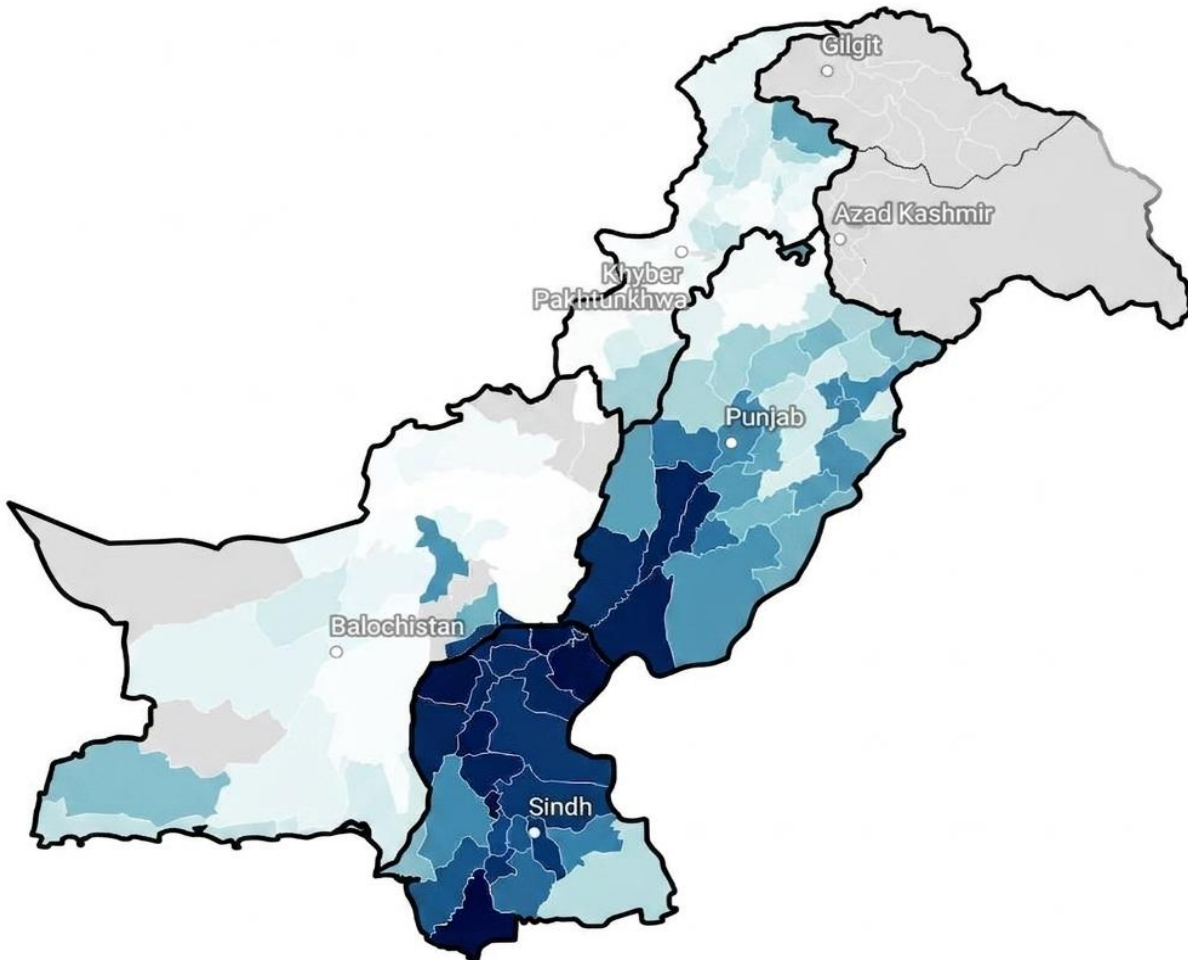
Figure 2 illustrates the spatial distribution of population flood exposure, multidimensional structural vulnerability, and district poverty rates across the 127 study districts. **Panel (a)** shows that population flood exposure was highest in the eastern and southern districts of Punjab and Sindh, particularly along the Indus River system, where exposure reached up to 96.4% of the district population. In contrast, many districts in Balochistan and Khyber Pakhtunkhwa exhibited low or no population flood exposure. **Panel (b)** shows that higher structural vulnerability scores were concentrated across much of Balochistan and northern Khyber Pakhtunkhwa, whereas many districts in Punjab exhibited comparatively lower levels of structural vulnerability. **Panel (c)** shows that higher poverty rates were concentrated primarily in Balochistan and interior Sindh.

Pakistan District wise Flood Exposure (%)

Flood exposure

0

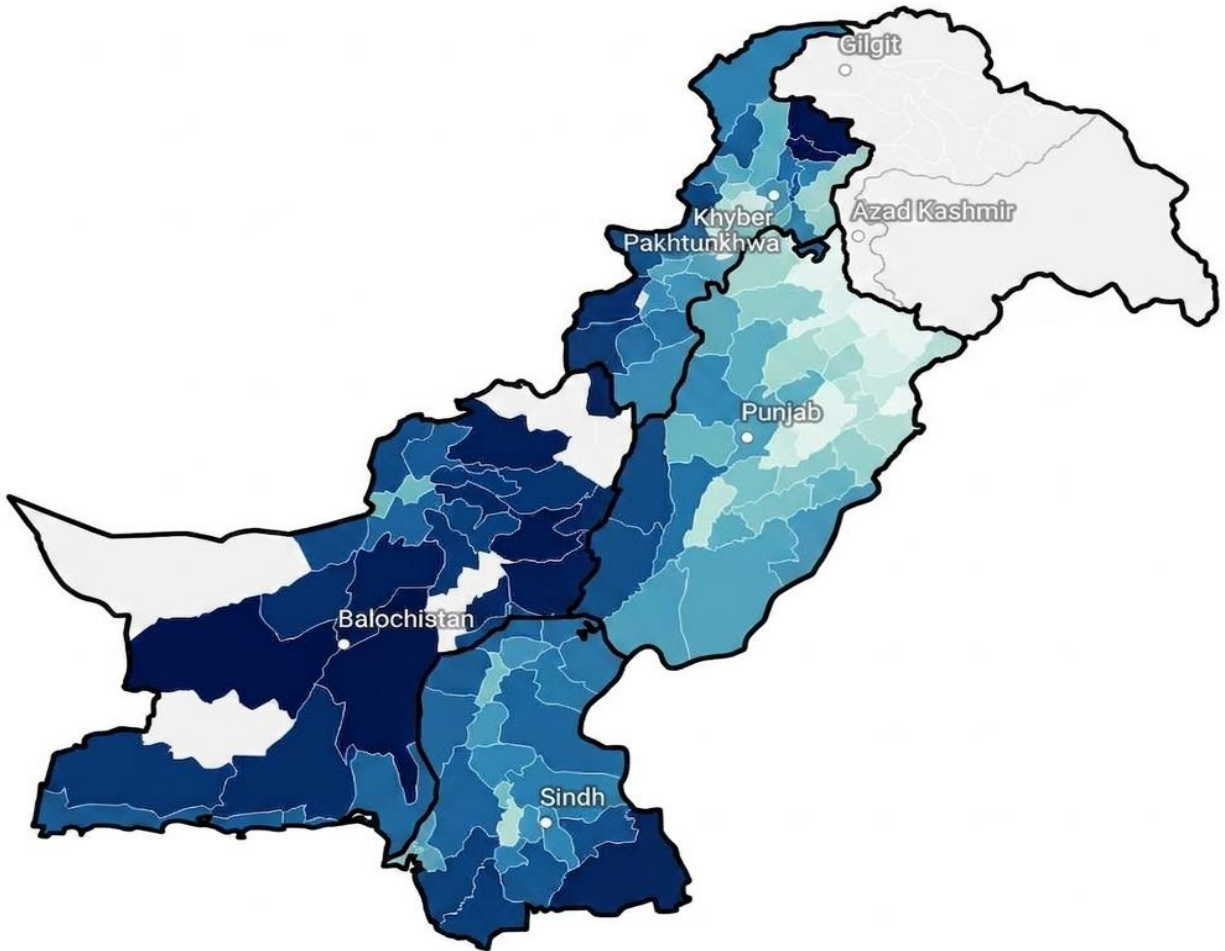
96.39



Pakistan District wise Vulnerability Index

Vulnerability Index

0 1



Pakistan District wise Poverty Rates

Poverty Rates

0.04

0.76

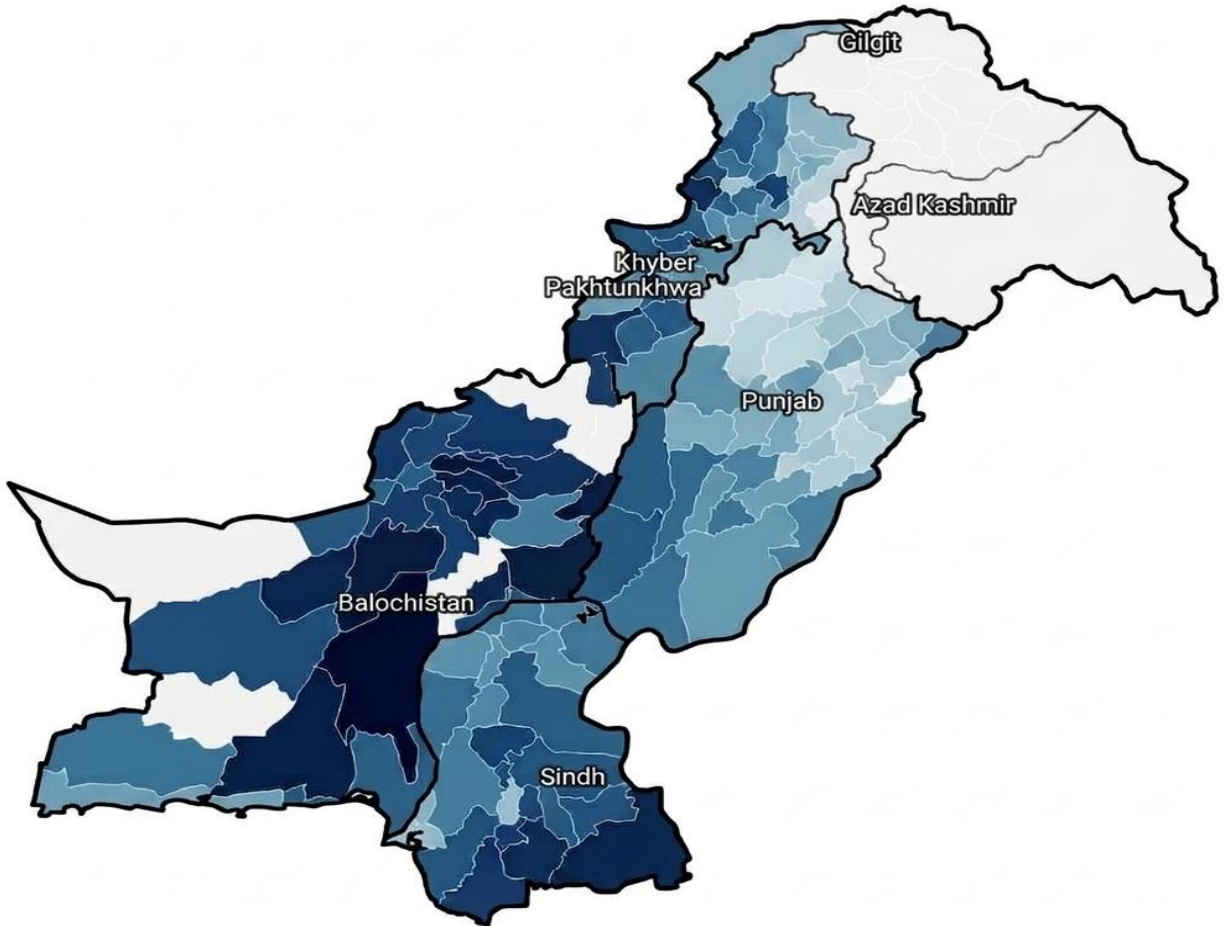


Fig. 2. Spatial distribution of population flood exposure, structural vulnerability, and district poverty rates across Pakistan. Panel (a) shows population flood exposure, (b) shows multidimensional structural vulnerability, and (c) shows district poverty rates. Note: The poverty rate is presented for descriptive spatial comparison only and was not included in the regression or Shapley decomposition analyses, which use district-level consumption measures as the indicators of regional welfare.

3.3 Bivariate Relationships

Figure 3 presents the bivariate relationships between the two consumption measures and the principal explanatory variables. **Panel (a)** shows little association between mean district consumption and population flood exposure, with a near-horizontal fitted line and an R^2 of 0.005, indicating that population flood exposure alone explained less than 1% of the variation in consumption. In contrast, **panel (b)** shows a markedly stronger relationship between mean district consumption and the structural vulnerability index, with an R^2 of 0.626. The same overall pattern was observed for consumption among poor households (**panels (c) and (d)**). Population flood exposure showed little association with consumption among poor

households, whereas structural vulnerability exhibited a stronger negative relationship, although the explained variation ($R^2 = 0.329$) was lower than that observed for mean district consumption.

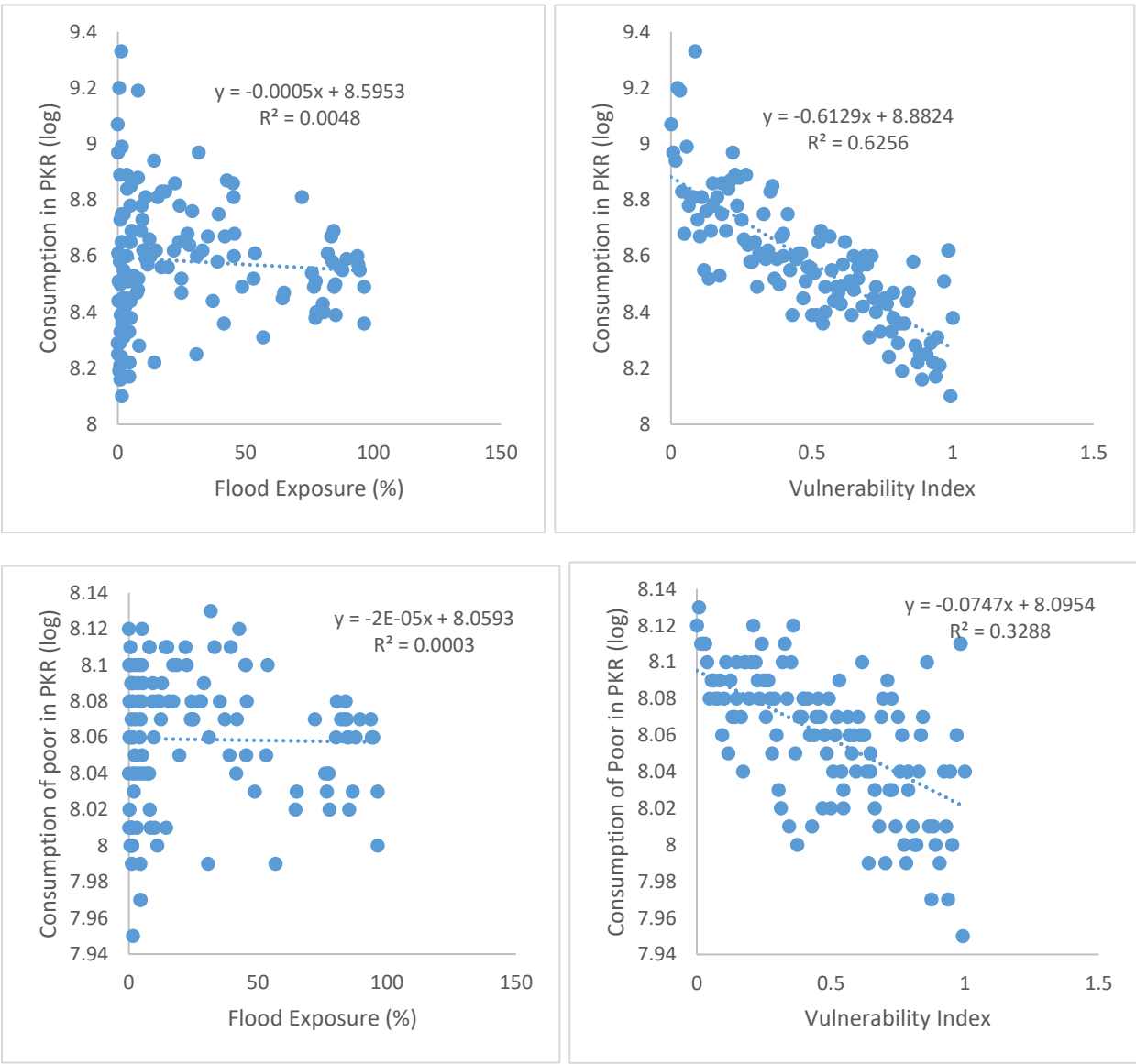


Fig. 3. Bivariate relationships between population flood exposure, structural vulnerability, and district-level consumption outcomes. **Panels (a) and (b)** show mean district consumption per adult equivalent, while **panels (c) and (d)** show mean consumption among poor households.

3.4 Correlation Analysis

Table 3 presents the Pearson correlation matrix for the study variables. Mean consumption was negatively correlated with all six structural vulnerability domains, with the strongest correlations observed for the Livelihoods ($r = -0.763$, $p < 0.001$) and Housing ($r = -0.752$, $p < 0.001$) domains. Similar negative correlations were observed for consumption among poor, although the magnitudes were generally smaller. Population flood exposure was not significantly correlated with either consumption outcome. In contrast, temperature change was negatively correlated with both mean consumption ($r = -0.379$, $p < 0.001$) and

consumption among the poor ($r = -0.452$, $p < 0.001$), whereas rainfall change showed no significant correlation with either outcome. The six structural vulnerability domains were positively correlated with one another, reflecting the multidimensional nature of structural vulnerability. These relationships were considered in the subsequent multivariable analysis through the use of the composite District Vulnerability Index.

Table 3 Pearson correlation matrix of study variables.

Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Consumption	1.000										
Consumption of the Poor	0.811** *	1.000									
Population Flood Exposure	-0.070	-0.009	1.000								
Housing	-0.752** *	-0.561** *	0.193*	1.000							
Communication & Transport	-0.692** *	-0.509** *	0.053	0.774** **	1.000						
Livelihoods	-0.763** *	-0.534** *	0.276* *	0.836* **	0.702* **	1.000					
Health	-0.666** *	-0.500** *	-0.295* **	0.709* **	0.805* **	0.601* **	1.000				
Education	-0.640** *	-0.469** *	-0.073	0.711* **	0.837* **	0.707* **	0.863 ***	1.000			
Demographics	-0.713** *	-0.510** *	0.047	0.787* **	0.581* **	0.768* **	0.606 ***	0.576 ***	1.000		
Temperature Change	-0.379** *	-0.452** *	0.310* **	0.401* **	0.416* **	0.443* **	0.286 **	0.442 ***	0.232 **	1.000	
Rainfall Change	0.060	0.087	-0.456* **	-0.182*	-0.177*	-0.268* *	0.048	-0.219 *	-0.016	-0.730* **	1.000

Note: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

3.5 Regression Analysis

3.5.1 Baseline Models: Climate and Flood Exposure Predictors Only (Models A1 and B1)

Table 4 presents OLS regression results with Conley spatial standard errors, which account for spatial dependence between geographically proximate districts. In the baseline models (A1 and B1), higher population flood exposure was negatively associated with both welfare outcomes. Rainfall change was also negatively associated with welfare, while temperature change was not statistically significant for mean consumption (Model A1) but was negatively associated with consumption among poor households (Model

B1). The baseline models explained 46.1% of the variation in mean consumption and 40.1% of the variation in consumption among poor households (adjusted $R^2 = 0.461$ and 0.401 , respectively). Relative to Punjab, Khyber Pakhtunkhwa and Balochistan had significantly lower mean consumption in Model A1, whereas the coefficient for Sindh was not statistically significant. For consumption among poor households, the province coefficients were comparatively smaller, with Khyber Pakhtunkhwa and Balochistan showing significantly lower consumption than Punjab.

3.5.2 Full Models: Introducing Structural Vulnerability (Models A2 and B2)

Adding the composite structural vulnerability index substantially improved model fit (Table 4). In Model A2, the vulnerability index was strongly and negatively associated with mean consumption ($\beta = -0.519$, $p < 0.001$). The adjusted R^2 increased from 0.461 in the baseline model to 0.665 in the full specification. Population flood exposure remained negatively associated with mean consumption and remained statistically significant after accounting for structural vulnerability. Temperature change was not statistically significant, while rainfall change was negative but did not reach the 5% significance threshold under the Conley spatial HAC specification. The province coefficients were no longer statistically significant after the inclusion of structural vulnerability. For consumption among poor households (Model B2), the vulnerability index was also negatively associated with the outcome ($\beta = -0.048$, $p < 0.01$). The adjusted R^2 increased from 0.401 to 0.461 . In contrast to Model A2, population flood exposure was not statistically significant after structural vulnerability was introduced. Temperature change remained negatively associated with consumption among poor households ($p < 0.05$), whereas rainfall change was not statistically significant. None of the province coefficients were statistically significant in the full specification.

Table 4 OLS Regression results with Robust Standard Errors

Variables	Average consumption		Consumption of the poor	
	Model A1 (Baseline)	Model A2 (Full)	Model B1 (Baseline)	Model B2 (Full)
Flood exposure	-0.003^{**} (0.001)	-0.001^* (0.001)	-0.000 (0.000)	-0.000 (0.000)
Vulnerability index	—	-0.519^{***} (0.077)	—	-0.048^{**} (0.017)
Temperature change	-0.067 (0.094)	-0.092 (0.059)	-0.034^* (0.016)	-0.037^* (0.018)
Rainfall change	-0.001 (0.001)	-0.001 (0.000)	-0.000 (0.000)	-0.000 (0.000)
Province dummies (ref: Punjab)				
Sindh	-0.056 (0.082)	0.033 (0.042)	-0.013 (0.014)	-0.004 (0.012)
Khyber Pakhtunkhwa	-0.180^{**}	-0.038	-0.027^*	-0.014

	(0.055)	(0.054)	(0.011)	(0.013)
Balochistan	−0.442***	−0.094	−0.056***	−0.024
	(0.081)	(0.058)	(0.012)	(0.015)
Constant	8.897***	8.976***	8.114***	8.121***
	(0.059)	(0.046)	(0.012)	(0.013)
Model fit				
R²	0.461	0.665	0.401	0.461
Observations	127	127	127	127
Province fixed effects	Yes	Yes	Yes	Yes
Conley cutoff (km)	175	175	175	175
Mean neighbours within cutoff	15.8	15.8	15.8	15.8

Notes: Conley spatial heteroskedasticity and autocorrelation consistent (HAC) standard errors are reported in parentheses. The distance metric is the haversine distance between district centroids, with a 175-km cutoff. Punjab is the reference province. * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

3.6 Shapley Value Decomposition: Domain-Level Relative Importance

Table 5 presents the Shapley decomposition of the predictors' relative contributions to the explained variation in the two welfare outcomes. The decomposition model explained 70.6% of the variation in mean district consumption and 48.8% of the variation in consumption among poor. For mean consumption, the largest contributions were from the structural vulnerability domains, particularly livelihoods (21.2%), housing (16.9%), and demographics (16.7%). Flood exposure contributed 0.9% of the explained variation. For consumption among poor, temperature change was the largest contributor (25.6%), followed by housing (14.3%), livelihoods (12.5%), and demographics (12.1%). Flood exposure accounted for 1.4% of the explained variation.

Table 5 Relative Importance of Predictors for District Consumption Outcomes

Predictor	Mean Consumption (lmg)	% R ²	Consumption Poor (lmg)	of % R ²
Livelihoods	0.150	21.2	0.061	12.5
Housing	0.119	16.9	0.070	14.3
Demographics	0.118	16.7	0.059	12.1

Communication & Transport	0.095	13.5	0.048	9.8
Health	0.089	12.6	0.050	10.2
Education	0.072	10.2	0.038	7.8
Temperature Change	0.040	5.7	0.125	25.6
Rainfall Change	0.017	2.4	0.029	5.9
Flood Exposure	0.006	0.9	0.007	1.4
Model R ²	0.706	100.0	0.488	100.0

4. Discussion

This study examines the geography of population flood exposure, structural vulnerability, and welfare across 127 districts of Pakistan and contributes to the literature in three complementary ways. First, welfare is assessed using two consumption-based measures—average district consumption and consumption among poor households—allowing the spatial correlates of overall welfare to be distinguished from those associated with poorer populations. Second, the analysis uses a population-based measure of flood exposure, defined as the proportion of district residents located within a modelled 100-year flood inundation footprint, rather than an area-based or purely physical measure of flood hazard. Third, and most centrally, the study compares the relative explanatory importance of population flood exposure and multidimensional structural vulnerability within a common geographic framework. Using the nationally representative and externally validated District Vulnerability Index for Pakistan (DVIP) provides a consistent basis for examining how these dimensions are distributed across the national district landscape.

A global study of Rentschler et al. [6] showed that flood exposure and poverty are concentrated in the same regions, with high exposure and high poverty co-occurring in much of South Asia and Sub-Saharan Africa. Their analysis demonstrated the value of considering the characteristics of exposed populations alongside exposure itself. However, our subnational analysis for Pakistan reveals a different pattern: high population flood exposure and high structural vulnerability are not spatially coincident. This geographic shift allows the distribution of exposed populations to be considered together with the persistent characteristics of places and their relative explanatory importance for welfare differences. The results show that districts containing larger shares of potentially exposed populations are not necessarily those with higher structural vulnerability, revealing a geographic distinction that can be obscured by national-level aggregation.

Population flood exposure accounted for only a small share of the explained variation in both welfare measures, whereas the six structural vulnerability domains collectively accounted for a substantially larger share, particularly for average consumption. This pattern does not imply that population flood exposure is unimportant for welfare or disaster risk; rather, it indicates that the spatial distribution of exposed populations provides different information from the persistent structural characteristics of the places in which they live. In Pakistan, population flood exposure is concentrated largely along the Indus River system and across parts of Punjab and Sindh, whereas structural vulnerability is concentrated across much of Balochistan and northern Khyber Pakhtunkhwa. These patterns therefore do not closely coincide, consistent with the broader literature treating exposure and vulnerability as related but distinct dimensions of environmental risk [9–12].

The stronger contribution of structural vulnerability can be understood in terms of the persistent characteristics of places. Housing, livelihoods, access to services, and demographic composition reflect social, infrastructural, and demographic conditions that develop and persist over time and distinguish districts from one another. In this sense, the vulnerability measures capture a form of structural stock: relatively enduring characteristics of places that provide the context in which populations experience environmental risks. The six domains collectively accounted for 91.1% of the explained variation in average consumption, with livelihoods, housing, and demographics making the largest individual contributions. Livelihoods capture differences in economic opportunities and security; housing reflects material living conditions relevant to environmental stress; and demographic composition reflects differences in population structure and dependency [11,20,29,30]. Their prominence indicates that measurable differences in the structural characteristics of districts correspond closely to the spatial distribution of welfare.

The geographic patterning of structural vulnerability further reinforces this interpretation. Higher vulnerability is concentrated in particular regions of Pakistan rather than being evenly distributed across districts, while the vulnerability domains are themselves spatially and statistically interrelated. Neighbouring districts may therefore share similar structural and welfare characteristics because they are situated within common economic, infrastructural, demographic, and environmental contexts. This is supported by the spatial dependence diagnostics: once the vulnerability index was introduced, less geographic structure was left unexplained in the models for both welfare outcomes, suggesting that much of the underlying spatial pattern in welfare is already being captured by structural vulnerability itself. Although this does not identify a specific mechanism, it is consistent with the interpretation that measured structural vulnerability captures part of the geographic patterning of welfare across districts.

The comparison of the two welfare outcomes adds an important distributional dimension. For average consumption, structural vulnerability was the dominant source of explained variation, with livelihoods, housing, and demographics among the largest individual contributors. For consumption among poor households, however, long-term temperature change was the largest individual contributor, accounting for 25.6% of the explained variation, while the structural vulnerability domains collectively remained more important overall. This difference indicates that the spatial correlates of welfare are not uniform across welfare groups. Poorer populations may be more sensitive to gradual climatic change because of greater reliance on climate-sensitive livelihoods and more limited resources for absorbing persistent environmental pressures [11–17]. Examining both welfare outcomes therefore reveals a distributional pattern that would be less visible from average welfare alone: structural characteristics remain central across the two outcomes, while the relative contribution of climatic change becomes more prominent for poorer populations.

Research on floods and vulnerability in Pakistan has largely examined individual flood events, specific provinces or communities, or individual dimensions of vulnerability [20,26–30]; this study provides a broader subnational geographic perspective. By combining a nationally consistent population exposure measure with the DVIP and two complementary welfare outcomes, it shows that the geography of population exposure and the geography of structural vulnerability are related but distinct, and that their relative explanatory importance for welfare also differs. The contribution is therefore not simply to identify where exposure or vulnerability is high, but to examine how these geographically differentiated characteristics correspond to spatial differences in welfare across the country.

These distinctions have implications for spatially targeted adaptation and disaster-risk planning in Pakistan, and may extend to other flood-exposed countries where structural vulnerability varies substantially across subnational units. Population flood exposure remains relevant for identifying where larger shares of the population are potentially exposed under a severe flood scenario, but exposure information alone may provide an incomplete basis for identifying districts with greater welfare disadvantage. Because districts with higher population flood exposure are not necessarily those with higher structural vulnerability,

considering exposure alongside livelihoods, housing, demographic conditions, and other dimensions of structural vulnerability could provide a more differentiated basis for adaptation and development planning [11,18,19].

5. Limitations and Future Research

Several limitations should be considered when interpreting these findings. First, the analysis is based on cross-sectional district-level data and therefore identifies statistical associations rather than causal relationships. The study was designed to examine long-term spatial differences in regional welfare rather than the occurrence, frequency, or severity of individual flood events. Accordingly, the explanatory variables represent relatively stable district characteristics: population flood exposure reflects a standardized 100-year flood inundation footprint, the climate variables capture long-term changes in temperature and precipitation, and the multidimensional vulnerability index measures enduring structural conditions. This focus on persistent characteristics provides an appropriate basis for comparing regional welfare across districts, although causal inference is beyond the scope of the analysis.

Second, the explanatory variables were compiled from nationally representative datasets collected over different, although partially overlapping, time periods. Because the selected indicators describe relatively stable geographic, climatic, and structural characteristics that generally change gradually over time, these temporal differences are unlikely to substantially affect the observed district-level associations. Moreover, harmonizing nationally representative datasets collected in the same year is rarely feasible in developing-country settings [38], making the integration of the most recent available datasets a practical approach for regional cross-sectional analysis. Nevertheless, some residual measurement error arising from temporal mismatches cannot be excluded.

Third, the district-level design cannot capture important within-district heterogeneity in household characteristics, including asset ownership, livelihood diversification, informal support networks, and adaptive capacity. In addition, the central objective of this study was to apply Shapley value decomposition to quantify the relative explanatory importance of the explanatory variables, an approach that is most readily implemented within the ordinary least squares (OLS) framework. Although province fixed effects account for broad regional heterogeneity, the models do not explicitly account for spatial dependence between neighbouring districts. Future research could combine household-level panel data with district-level structural indicators or apply spatial econometric approaches to examine the robustness of these findings while accounting for spatial dependence. Furthermore, studies employing longitudinal data or stronger causal identification strategies would help clarify the causal pathways linking flood exposure, vulnerability, and regional welfare.

References

1. Intergovernmental Panel on Climate Change. *Climate change 2023: Synthesis report*. IPCC (2023).
2. United Nations Office for Disaster Risk Reduction. *Global assessment report on disaster risk reduction 2024*. United Nations (2024).
3. Hirabayashi, Y. *et al.* Global flood risk under climate change. *Nat. Clim. Change* **3**, 816–821 (2013).
4. Winsemius, H. C. *et al.* Disaster risk, climate change, and poverty: Assessing the global exposure of poor people to floods and droughts. *Policy Res. Work. Pap.* **7480** (World Bank, 2015).
5. Dullaart, J. C. M. *et al.* Accounting for tropical cyclones more than doubles the global population exposed to low-probability coastal flooding. *Commun. Earth Environ.* **2**, 135 (2021).
6. Rentschler, J., Salhab, M. & Jafino, B. A. Flood exposure and poverty in 188 countries. *Nat. Commun.* **13**, 3527 (2022).

7. Kumar, D. & Bhattacharjya, R. K. Review of different methods and techniques used for flood vulnerability analysis. *Nat. Hazards Earth Syst. Sci.* **20**, 297–308 (2020).
8. Neal, J. & Hawker, L. A vulnerability-adjusted approach to global flood risk mapping. In *EGU General Assembly 2023* (2023).
9. Wisner, B., Blaikie, P., Cannon, T. & Davis, I. *At Risk: Natural Hazards, People's Vulnerability and Disasters*. 2nd edn. (Routledge, 2004).
10. Cutter, S. L., Boruff, B. J. & Shirley, W. L. Social vulnerability to environmental hazards. *Soc. Sci. Q.* **84**, 242–261 (2003).
11. Hallegatte, S., Bangalore, M., Vogt-Schilb, A. & Rozenberg, J. *Unbreakable: Building the Resilience of the Poor in the Face of Natural Disasters*. (World Bank, 2017).
12. Hallegatte, S., Vogt-Schilb, A., Rozenberg, J. *et al.* From poverty to disaster and back: A review of the literature. *Econ. Dis. Clim. Change* **4**, 223–247 (2020).
13. Dell, M., Jones, B. F. & Olken, B. A. Temperature shocks and economic growth: Evidence from the last half century. *Am. Econ. J. Macroecon.* **4**, 66–95 (2012).
14. Diffenbaugh, N. S. & Burke, M. Global warming has increased global economic inequality. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 9808–9813 (2019).
15. Azzarri, C. & Signorelli, S. Climate and poverty in Africa south of the Sahara. *World Dev.* **125**, 104691 (2020).
16. Mudzingiri, C., Mudiriza, G., Jana, G. & Sunge, R. The effect of local climatic conditions on household consumption: A case study of South Africa. *Humanit. Soc. Sci. Commun.* **11**, 652 (2024).
17. Carter, M. R. *et al.* Poverty traps and natural disasters in Ethiopia and Honduras. *World Dev.* **35**, 835–856 (2007).
18. Formetta, G. & Feyen, L. Empirical evidence of declining global vulnerability to climate-related hazards. *Glob. Environ. Change* **57**, 101920 (2019).
19. Erman, A. *et al.* *The Road to Recovery: The Role of Poverty in the Exposure, Vulnerability and Resilience to Floods in Accra. Policy Res. Work. Pap.* **8469** (World Bank, 2018).
20. Ahmad, I. A. & Khan, A. N. Climate change vulnerability and multidimensional poverty in flood-prone rural areas of Punjab, Pakistan: An application of the multidimensional poverty index and livelihood vulnerability index. *Environ. Dev. Sustain.* **25**, 14145–14169 (2023).
21. Germanwatch. *Global Climate Risk Index 2025*. (Germanwatch e.V., 2025).
22. Otto, F. E. L. *et al.* Climate change increased extreme monsoon rainfall, flooding highly vulnerable communities in Pakistan. *Environ. Res. Clim.* **2**, 025001 (2023).
23. Knippenberg, E., Jensen, N. & Conostas, M. Poverty impacts of the Pakistan flood 2022. *Econ. Dis. Clim. Change* **8**, 453–471 (2024).
24. Waseem, H. B. & Rana, I. A. Floods in Pakistan: A state-of-the-art review. *Nat. Hazards Res.* **3**, 359–373 (2023).
25. Government of Pakistan. *Fiscal Risk Statement 2024–25*. Ministry of Finance (2025).
26. Kurosaki, T. Vulnerability of household consumption to floods and droughts in developing countries: Evidence from Pakistan. *Environ. Dev. Econ.* **20**, 209–235 (2015).
27. Shah, A. A., Ye, J., Abid, M., Khan, J. & Amir, S. M. Flood hazards: Household vulnerability and resilience in disaster-prone districts of Khyber Pakhtunkhwa Province, Pakistan. *Nat. Hazards* **93**, 147–165 (2018).
28. Manzoor, W. Analysis of the vulnerability of farm households to flood risk in Punjab, Pakistan. *J. Econ. Impact* **3**, 27–38 (2021).
29. Hamidi, A. R. *et al.* Flood exposure and social vulnerability analysis in rural areas of developing countries: An empirical study of Charsadda District, Pakistan. *Water* **14**, 1176 (2022).
30. Rasool, S., Rana, I. A. & Waseem, H. B. Assessing multidimensional vulnerability of rural areas to flooding: An index-based approach. *Int. J. Disaster Risk Sci.* **15**, 88–106 (2024).
31. United Nations Development Programme Pakistan. *Pakistan and Climate Change: Vulnerability and Response*. United Nations Development Programme (2023).

32. Population Council Pakistan. *District Vulnerability Index for Pakistan: Harnessing Multisectoral Data to Inform Equitable Policy and Climate Action*. Population Council Pakistan (2025).
33. Garza-Rodriguez, J. Poverty and economic growth in Mexico. *Soc. Sci.* **7**, 183 (2018).
34. Deaton, A. & Zaidi, S. *Guidelines for Constructing Consumption Aggregates for Welfare Analysis. LSMS Working Paper 135* (World Bank, 2002).
35. Haughton, J. & Khandker, S. R. *Handbook on Poverty and Inequality*. (World Bank, 2009).
36. Conley, T. G. GMM estimation with cross-sectional dependence. *J. Econom.* **92**, 1–45 (1999).
37. Lindeman, R. H., Merenda, P. F. & Gold, R. Z. *Introduction to Bivariate and Multivariate Analysis*. (Scott, Foresman, 1980).
38. Grömping, U. Estimators of relative importance in linear regression based on variance decomposition. *Am. Stat.* **61**, 139–147 (2007).
39. **United Nations**. *Report of the Secretary-General on Household Surveys*. Statistical Commission, Forty-sixth Session (E/CN.3/2015/10) (2015). Available at: <https://unstats.un.org/unsd/statcom/doc15/2015-10-HouseholdSurveys-E.pdf>

Author Contributions: Conceptualization, methodology, data curation, formal analysis, investigation, visualization, writing original draft preparation, review and editing, and final approval of the manuscript. The author is solely responsible for the conception, design, analysis, interpretation, and reporting of the study.

Data Availability Statement

All datasets used in this study are publicly available from the following repositories:

World Bank Pakistan Poverty Mapping 2019–2020: <https://datacatalog.worldbank.org/search/dataset/0066822/pakistan-poverty-map-2019-2020>
Humanitarian Data Exchange (HDX) Pakistan Risk Assessment Indicators: <https://data.humdata.org/dataset/pakistan---risk-assessment-indicators>
District Vulnerability Index for Pakistan (DVIP): <https://dvp.pcpak.org/DVIP-Report-2025.pdf>

Funding Statement

No funding was received for the research reported in this article.

Disclosure Statement

The author reports there are no competing interests to declare.

Declaration of Generative AI Use

The author used AI tools to assist with language editing, manuscript organisation, and improving the clarity of the writing. All research design, statistical analyses, interpretation of results, and final editorial decisions were undertaken and verified by the author, who takes full responsibility for the accuracy, integrity, and originality of the manuscript.

Ethics Statement

Ethical approval was not required for this study because it is based entirely on publicly available secondary data and did not involve human participants or identifiable personal information.