The burden of flood risk

Fiona Burlig, Jack N. Glaser, and Anant Sudarshan*

September 2026

Abstract

Floods are enormously harmful, but their economic costs may extend well beyond direct damages. We develop a method for measuring the economic costs of flood risk, combining both damages from realized floods and the persistent return gap from living in a high-risk environment. We implement it among smallholder farmers in Odisha, India. Otherwise similar high-risk villages have 43% lower expected returns than nearby low-risk villages. Expected losses from realized floods account for only roughly half of this gap; the remainder comes from reduced returns across all seasons. Focusing on realized damages alone substantially understates the economic burden of flood risk.

*Burlig: Harris School of Public Policy, University of Chicago and NBER. Email: burlig@uchicago.edu. Glaser: Kenneth C. Griffin Department of Economics, University of Chicago. Email: glaserj@uchicago.edu. Sudarshan: Department of Economics, University of Warwick. Email: anant.sudarshan@warwick.ac.uk. We are grateful to Josh Dean and Erin Kelley for helpful comments and suggestions. We thank Nikilesh Anusha, Zoya Khan, Yashaswi Mohanty, Suyash Nandgaonkar, Kira Silvestrovich, and Camellia Ye for outstanding research assistance, and Urmila Chatterjee for helpful project management. We thank the Government of Odisha's Data, Policy, and Innovation Centre for project support. This research has IRB approval from the University of Chicago (Protocol No. IRB24-0685), and was funded by the Data, Policy, and Innovation Centre and the Becker Friedman Institute at the University of Chicago. All remaining errors are our own.

1 Introduction

Almost two billion people – 89% of them in low- and middle-income countries – live in areas that are exposed to substantial flood risk (Rentschler et al., 2022). Climate change is projected to sharply intensify this hazard, with increases concentrated in the developing world (Hirabayashi et al., 2021; IPCC, 2022).[1] Yet, we have little empirical evidence quantifying the full costs of flood risk on the rural poor. Although the most visible costs of flooding are damages directly caused by inundation, the economic burden of flood *risk* extends much further. This is because households and firms who make investment and production choices with future flooding in mind may sacrifice returns in ordinary seasons to reduce losses when floods occur. Consequently, focusing on direct damages alone may substantially understate the burden of flood risk, leading policymakers to undervalue the benefits of risk reduction relative to disaster relief.

In this paper, we provide an estimate of the full costs of flood risk on smallholder farmers in flood-prone Odisha, one of India's most climate-vulnerable states. To do so, we develop a simple accounting framework in Section 2 that separates the burden of flood risk into two pieces: the all-season component, incurred in every season, regardless of whether a flood takes place (e.g., purchasing flood-tolerant seeds), and the realized-flood component, which accounts for the probability of a flood occurring and the resulting damages (e.g., crop revenue lost to inundation). We then design a sampling frame for a bespoke primary survey that enables credible identification of the two components; this approach is designed to be useful in settings beyond our own.

Empirically estimating the burden of living with flood risk requires two distinct types of identifying variation. As the all-season component is an equilibrium response to a persistent risk environment, it requires cross-sectional variation in flood risk that allows the econometrician to compare otherwise similar locations with different likelihoods of inundation. In contrast, the realized-flood component can be estimated using year-to-year variation in realized flooding within a given risk environment. Existing work on floods has largely relied on the latter, comparing outcomes before and after floods or across flooded and unaffected locations to estimate the damages from realized flooding (Kocornik-Mina et al., 2020; Patel, 2024; Balboni et al., 2026).[2] However, this variation

[1]The population exposed to river flooding is expected to more than double under a 2°C warming scenario, and increase roughly five-fold under a 4°C warming scenario (Caretta et al., 2022).

[2]A related literature has used randomized trials to evaluate adaptation technologies (Emerick et al., 2016; Lane, 2024), but similarly does not measure the overall burden of flood risk.

alone cannot identify the effects of persistent differences in flood risk on economic outcomes. In addition, measuring the economic burden of flood risk on farmers requires a welfare measure such as profits or utility, on which secondary data are very limited.[3]

Our first step is to identify plausibly exogenous cross-sectional variation in the risk environment, which we use to select our survey sample. We take advantage of a common characteristic of flood-prone regions that is well-documented in the hydrology literature: highly localized variation in topography leads to substantively different probabilities of flooding between two villages, even within small geographic areas (Choi and Harvey, 2014; Bates et al., 2006; Scown et al., 2015; Merz et al., 2021). We use Odisha's Flood Hazard Atlas (National Remote Sensing Centre, 2019) to identify a sample of farmers who live close together but face very different levels of flood risk. We conducted a bespoke survey of 2,141 households in 401 villages in two flood-prone districts of the state, stratified on administrative blocks, each roughly the size of a US ZIP code, with an average radius of 9.2 km. This allows us to compare villages that face the same administrative policies, similar market access, and similar infrastructure, but differ in flood probability due to highly localized topographical features. To generate plausibly exogenous variation in idiosyncratic flooding, we conduct this survey in both the 2023 and 2025 kharif agricultural seasons, allowing us to observe farmers in both ordinary and flood conditions. We collect panel data on farm revenues, costs, and investment decisions, allowing us to construct a measure of net agricultural returns. We describe our empirical approach in Section 3.

We combine our cross-sectional variation in flood risk with idiosyncratic variation in realized flooding to estimate three empirical objects in a single regression: (1) the effect of flood risk on net agricultural returns in ordinary seasons, which pins down the all-season component of the burden; (2) the effect of a realized flood in a low-risk environment; and (3) the additional effect of a realized flood in a high-risk environment. Together with the historical probability of flooding in each environment, measured using the Flood Hazard Atlas, the latter two objects pin down the realized-flood component. The sum of the two components is the overall annual burden of flood risk. We present the results in Section 4.

[3] This is in contrast to data on crop yields where rich spatial information is often available in administrative data and through remote sensing. The environmental economics literature has accordingly used yields to study the impacts of climate change on agriculture (Hultgren et al., 2025; Burke and Emerick, 2016; Schlenker and Roberts, 2009). However, since farmers maximize profits – not yields – it is not straightforward to interpret such estimates as a measure of economic damages (Fisher et al., 2012; Hsiang, 2016).

The all-season component is substantial. Farmers in high-risk villages earn 4,724 Rs. per household-year less than farmers in low-risk villages ($p = 0.040$), equivalent to 21% of the ordinary-season mean returns of 22,044 Rs. in low-risk villages. Consistent with this, we find evidence that flood risk affects production choices in ordinary seasons: farmers in high-risk villages are less likely to cultivate high-yield-variety rice, choosing instead to plant flood-tolerant or traditional seed. They also engage more in direct seeding, and apply less irrigation. They cultivate more land, partly offsetting the fact that per-acre returns are substantially lower in high-risk villages.

Perhaps unsurprisingly, floods themselves reduce net returns. The point estimates are large: 5,962 Rs. (27% of the observed low-risk ordinary-season mean, $p = 0.004$) and 8,571 Rs. (42% of the observed high-risk ordinary-season mean, $p < 0.001$) for farmers in low- and high-risk villages, respectively, but are not distinguishable from each other ($p = 0.34$ on the difference).

The full burden of living with high flood risk is thus the all-season component *plus* the additional damage from a flood, weighted by the probability of flooding. In our area, high-risk villages are far more likely to see a flood than low-risk villages (57 vs. 3% of years), making the expected effects of realized flooding worse even absent differences in realized damages. The total expected annual burden of high flood risk is 9,415 Rs. per household (approximately 114 USD), or 43% of low-risk returns. 50% of this burden arises from the all-season component and 50% from the realized-flood component. Thus, focusing only on damages that result from realized floods will miss roughly half the total economic burden of high flood risk.

Our main contribution is to measure the full economic burden of persistent flood risk in agriculture by quantifying both the effects that are borne every season and the impacts of realized flooding. We build on two literatures. The first is previous empirical work estimating the costs of *realized* floods on households, firms, cities, migration, and agricultural production (Patel, 2024; Kocornik-Mina et al., 2020; Hornbeck and Naidu, 2014; del Ninno et al., 2003; Balboni et al., 2026). The second is an older and broader literature on agricultural insurance and risk which has shown that exposure to risk changes investment, technology adoption, and input choices *before* shocks occur – changing outcomes in all states of the world (Rosenzweig and Binswanger, 1993; Karlan et al., 2014; Mobarak and Rosenzweig, 2013; Donovan, 2021).[4] We connect these two margins in

[4]Our mechanism requires neither risk aversion nor uninsured risk. Even a risk-neutral firm may invest in costly adaptation that lowers returns in ordinary years but raises them in flood years; as flood risk rises, the expected-return-maximizing production bundle changes.

a common empirical accounting, estimated within the same population. This is similar in spirit to Jia et al. (2026), who use a spatial model to document the substantial role of long-run adjustments to flood risk by firms and workers in the United States. We focus on the very different setting of smallholder farmers in a developing country. We use a reduced-form approach with primary data, combined with a sampling strategy that delivers the identifying variation we need to measure both components of the overall burden. Our finding that both components are comparable, and large, underscores why joint measurement is important.

Our second contribution is methodological: we provide a strategy for directly measuring the equilibrium burden of flood risk in a developing-country agrarian setting. This augments two existing approaches in the literature. The first uses land prices as a summary statistic that incorporates all costs of risk (Mendelsohn et al., 1994; Hornbeck, 2012). Unfortunately, this is infeasible in many developing-country settings where land markets are thin or absent (Fernando, 2022; Acampora et al., 2025). The second uses responses to fluctuations in weather, combined with additional assumptions about optimizing behavior, to infer the costs of climate change (risk) net of adaptation (Deryugina and Hsiang, 2017; Hsiang, 2016; Carleton et al., 2022). This requires spatial data on outcomes and climate that vary smoothly over the distributions of interest; this is rarely the case for floods in developing countries.

In our complementary approach, we instead combine tightly localized cross-sectional variation in persistent flood risk, idiosyncratic realized floods, historical inundation frequencies, and a bespoke survey which measures net returns to directly estimate both the all-season and the realized-flood components of the overall burden of living with flood risk. Because these ingredients exist, or can be generated, in many flood-prone contexts, this combination of hazard histories, targeted spatial sampling, and rich economic measurement may provide a useful template for other settings.

Finally, our results have implications for a long-standing debate over the balance between financing disaster losses and risk reduction (Clarke and Dercon, 2016; Hallegatte et al., 2017; Kousky, 2019). Risk-transfer policies such as insurance or emergency loans shift resources toward flood states, which can improve welfare (Mobarak and Rosenzweig, 2013; Karlan et al., 2014; Lane, 2024). However, they do not eliminate flood damages, so even a risk-neutral farmer with full, actuarially fair insurance will continue to make costly adjustments that reduce flood exposure (i.e., our all-season component). Reducing these costs instead requires lowering expected damages, either by lowering the frequency of floods through infrastructure like dams or levees (Bradt and Aldy, 2025) or by decreasing damages conditional on flooding, for instance with advance warnings (Burlig et al., 2024; Jagnani and Pande, 2024) or flood-tolerant seeds (Emerick et al., 2016).[5] Our finding that roughly half of the burden of flood risk comes from the all-season component suggests that valuing flood policy using realized damages alone can significantly undervalue risk reduction relative to risk transfer.

2 Conceptual framework

Consider a farm operating under a Low or High risk environment, $r \in \{L, H\}$, where a damaging flood ($F \in \{0, 1\}$) occurs with probability p_r, with $p_H > p_L$. Before F is realized, the farmer chooses a production bundle z – seed varieties, fertilizer, irrigation, planting method, cultivated area, and other inputs. Profits are the state-contingent function

$$A(z, F \mid r).$$

The farmer chooses z to maximize expected profit:[6]

$$z_r \;\in\; \arg\max_{z}\, (1 - p_r)\, A(z, 0 \mid r) + p_r\, A(z, 1 \mid r).$$

Since z is contingent on risk, costs and revenues may differ in both flood seasons and ordinary seasons.

The all-season component Because the production bundle is chosen before the flood state is realized, the profit gap between environments has a component that is present in every season, flood or not, which we denote C_N. In an ordinary season, neither environment faces flood losses,

[5]Though it does not eliminate the underlying flood losses, compensation in the event of a flood (Innes, 2003; Christian et al., 2025) is similar in spirit: if it is free and fully covers damages, it removes the private incentive to make defensive investments. In practice, however, such compensation schemes are rarely comprehensive (Government of India, Ministry of Agriculture and Farmers Welfare, 2015).

[6]We focus on farm profits rather than other dimensions, as we find that flood risk does not impact off-farm income (Figure 3) or migration (Appendix Table B.3).

so the observed profit gap is exactly:

$$C_N \ \equiv\ A(z_L, 0 \mid L) - A(z_H, 0 \mid H). \tag{1}$$

Adding and subtracting $A(z_H, 0 \mid L)$ splits this into a reallocation of z and a residual effect of the environment at a fixed bundle:

$$C_N \ = \ \underbrace{A(z_L, 0 \mid L) - A(z_H, 0 \mid L)}_{\text{reallocation of } z} \ + \ \underbrace{A(z_H, 0 \mid L) - A(z_H, 0 \mid H)}_{\text{residual effect of } r \text{ at } z_H}.$$

The first term is the all-season cost of using the high-risk input bundle rather than the low-risk one. Note that if the environments only differ by the probability of flooding, p_r, by revealed preference, $A(z_L, 0 \mid L) \geq A(z_H, 0 \mid L)$: since z_L is optimal at p_L and z_H at $p_H > p_L$, the high-flood-risk bundle must (weakly) sacrifice returns in the ordinary state in exchange for returns in the flood state. As a result, any change in input allocation due to flood risk alone must have a weakly positive cost. The second term captures what remains when farmers in both environments use the same bundle: any channel through which the environment affects profits other than through z.[7] Separating the two terms would require estimating a production function; we instead treat C_N as a single object throughout.

Flood season During a flood season, losses (expressed as a positive number) are given by D_r where

$$D_r \ \equiv\ A(z_r, 0 \mid r) - A(z_r, 1 \mid r). \tag{2}$$

The gap between profits in the two risk environments in a flood season is $C_N + (D_H - D_L)$: the same all-season component plus the difference in losses from the realized flood. C_N is therefore borne in every season; when taking this to the data, we identify it in ordinary seasons only because this allows us to observe it separately from flood losses.[8]

[7] When the risk environment impacts profits only through changes in the probability of flooding, such that $A(\cdot \mid L) = A(\cdot \mid H)$, this term is zero, and C_N is simply the cost of choosing z_H instead of z_L. If the risk environment also affects the profit function then the second term will not be zero. For example, this would occur if unobserved differences in flood severity also correlate with frequency, or if repeated flood losses leave farms poorer and less productive even holding inputs fixed.

[8] Note that Equation (2) does not impose that $D_H < D_L$, though this would likely hold if the choice of z is 'defensive' (i.e., costly every season and productive only when $F = 1$). However, any cumulative effects from being

Calculating total costs therefore requires measuring both all-season *and* flood-season effects. Let $E_r \equiv \mathbb{E}_F\big[A(z_r, F \mid r)\big]$ denote expected annual profit in environment r; since a flood occurs with probability p_r, $E_r = (1 - p_r)\, A(z_r, 0 \mid r) + p_r\, A(z_r, 1 \mid r) = A(z_r, 0 \mid r) - p_r D_r$. Therefore, the expected annual burden of living with flood risk (again expressed positively), $C \equiv E_L - E_H$, is

$$C = C_N \; + \; \underbrace{p_H D_H - p_L D_L}_{C_F}, \tag{3}$$

the all-season component (C_N) – incurred in every season, regardless of whether a flood is realized – plus the realized-flood component (C_F), the probability-weighted gap in flood losses, which accounts for both the differential likelihood of floods in high- and low-risk environments (p_L, p_H) and differential losses when a flood occurs.

Equation (3) nests the objects captured by most prior empirical work on flooding: the losses when floods occur, D_r (Patel, 2024; Balboni et al., 2026; Kocornik-Mina et al., 2020; Hornbeck and Naidu, 2014; del Ninno et al., 2003), or expected damages from realized floods, $p_r \cdot D_r$ (Merz et al., 2010; Ward et al., 2013; Hallegatte et al., 2017). However, the difference in expected realized-flood losses ($p_H D_H - p_L D_L$) equals the full burden only in the special case where production bundles do not respond to risk ($z_H = z_L$) and the environment affects profits only through the probability of flooding, so that $C_N = 0$. Whenever households adjust their production bundle in response to risk, both D_r and $p_r \cdot D_r$ understate the true burden of living with flood risk, as they ignore C_N.

3 Research design

To compute the burden of living with flood risk, we must estimate three objects: (i) C_N, the all-season component, identified in ordinary seasons; (ii) D_L, flood damages in low-risk areas; and (iii) D_H, flood damages in high-risk areas. We therefore require a *cross-sectional* research design that allows us to compare outcomes in locations with high vs. low flood risk that are otherwise similar. To measure flood damages in high- and low-risk environments, we require estimates of the effects of idiosyncratic flooding on returns. The main elements of our approach are new data collection,

repeatedly flooded that change the production function, or flood severity that varies with risk environment, will affect D_H. For this reason, interpreting differences in flood-year costs as evidence of the presence/absence of effective adaptation is potentially misleading.

a prospective sampling strategy delivering the variation we need, and a regression specification to estimate C_N, D_L, and D_H.

3.1 Data

To estimate the economic costs of flooding we need a measure of profits or returns to agriculture, as well as flood risk and flood incidence. We obtain these from different data sources.

Farmer surveys First, we develop a panel survey instrument that gathers detailed information on kharif-season agriculture, including farmer revenue, inputs, crop choices and yields, livestock use, farming practices, off-farm work, cultivated area, profits, and reported flood histories. We administered the survey in 2023 and 2025. Our main outcome of interest is net agricultural returns, defined as the value of production (evaluated at block-median prices) net of costs.

Flood Hazard Atlas We also consult the Odisha government's Flood Hazard Atlas (National Remote Sensing Centre, 2019), which categorizes each village by the number of years of flooding it faced from 2001 to 2018, as measured by satellite. The Atlas categorizes villages into five hazard buckets: very low (1 flood year from 2001 to 2018), low (2–4 flood years), moderate (5–6 flood years), high (7–9 flood years), and very high (10–14 flood years). Because we have data on the universe of villages from the Census of India, we classify villages that do not appear in the Atlas as having no flooding. We use this data source to (i) design the sampling frame, (ii) classify villages into high vs. low risk for the purposes of analysis, and (iii) calibrate the probability of flooding (p_L, p_H).

Other data sources In addition to these main sources, we collect village boundaries, village population data, and infrastructure information from the SHRUG (Asher et al., 2021). We collect data on rice suitability from GAEZ (Fischer et al., 2021); soil information from SoilGrids (Poggio et al., 2021), elevation and slope from the European Space Agency's Copernicus satellite (European Space Agency, 2022); hydrological information from Copernicus matched with a hydrology model (Lindsay, 2016); and compute distance to river using information from OpenStreetMap (OpenStreetMap contributors, 2024). We geospatially match these gridded datasets to our villages.

3.2 Sampling and identification

Measuring the long-run all-season cost of flood risk requires plausibly exogenous spatial variation in risk. Our prospective sampling design exploits the fact that the micro-topography of floodplains frequently creates highly granular spatial variation in the propensity of pieces of land to experience inundation (Choi and Harvey, 2014; Bates et al., 2006; Scown et al., 2015; Merz et al., 2021). Such hyper-local variation means that villages within small geographic areas can experience very different flood histories, even as they are comparable on other factors that may affect agricultural output (government infrastructure, irrigation, electricity supply, access to markets, demographic characteristics, input subsidy regime, and so on). Land markets in Odisha are heavily restricted for statutory reasons (Government of Odisha, 1960; NITI Aayog, 2016), so farmers are highly limited in their ability to sort across locations according to productivity (Fernando, 2022; Rosenzweig and Wolpin, 1985; Mearns and Sinha, 1999; Deininger et al., 2009).

According to the Flood Hazard Atlas, Odisha's Kendrapara and Jajpur districts have substantial localized heterogeneity in flood risk across villages. Within these two districts, we sampled the blocks with the most variation in flooding per the Atlas: Aul, Rajkanika, Marsaghai, and Rajnagar in Kendrapara district and Jajpur, Dharmasala, and Binjharpur in Jajpur district. Blocks are small administrative units, roughly the size of a typical US ZIP code. In our sample, the average radius is 9.2 km.

From these blocks, we select 400 villages on the basis of flood hazard, removing middle classifications to maximize power: 120 high- or very-high-hazard villages per district, and 80 no-flood or very-low-hazard villages per district.[9] We refer to these as "high-risk" and "low-risk" villages, respectively, for the remainder of this analysis. Finally, we randomly sample 5–7 households per village, using a random walk procedure. Households that did not engage in agriculture were excluded from the survey. Attrition between the two waves was moderate; 1,861 households appear in both waves.

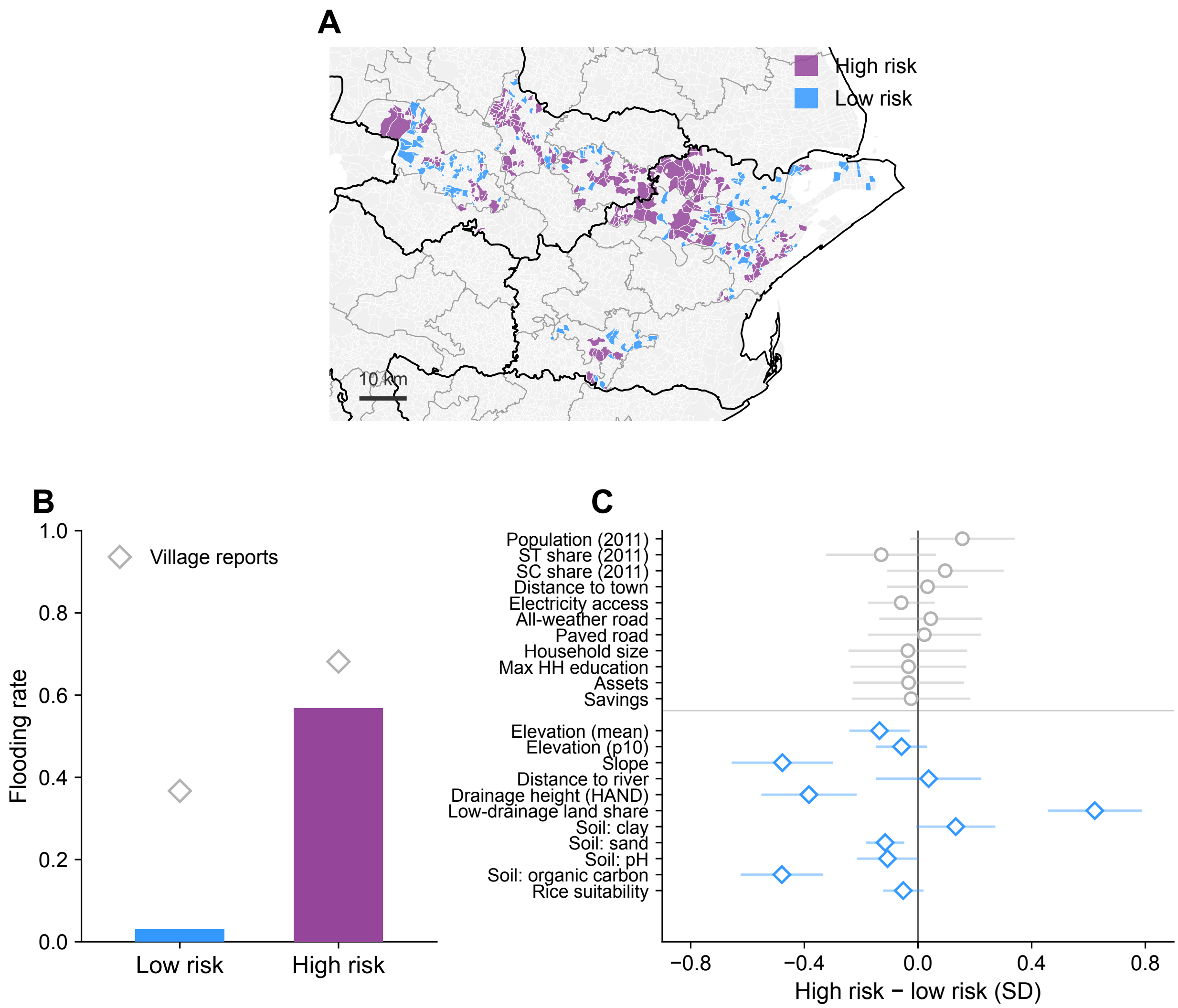
Figure 1 presents details about our selected villages, which support our research design. Appendix Table A.1 presents additional summary statistics. Panel A presents a map of our study area in the Kendrapara and Jajpur districts of Odisha. We plot low-risk villages in blue, and

[9]In the field, we ended up with one extra high-risk village, leaving our total sample size at 401.

high-risk villages in purple. This panel also shows blocks outlined in light gray. Panel B shows that the probability of flooding is substantially lower in low-risk villages than in high-risk villages, both measured using the Flood Hazard Atlas itself (bars) and using villagers' self-reported flood histories from our survey (diamonds). In our preferred accounting, we use the Atlas to establish the probability of flooding, because it provides a longer, independently measured historical record.[10] Finally, Panel C presents balance on observables between low- and high-risk villages after conditioning on block fixed effects. The top section (gray) shows that low- and high-risk villages are balanced on demographics, infrastructure, and wealth. The bottom section (blue) shows geological characteristics, several of which differ between low- and high-risk villages. This is essentially by construction: land slope, drainage height, and low-drainage land are what generate flood risk. Soil characteristics both contribute to flood risk and are changed by flooding. This figure demonstrates that while there is substantial variation in flood risk within our sampling frame, low- and high-risk villages are close together and balanced on a range of important characteristics.

[10]Instead using flood frequencies as reported in the survey does not meaningfully change our headline results; see Section 4.4.

Figure 1: Our research design compares similar villages with different flood risk



Notes: This figure plots the study design. Panel A maps the study-frame villages in Jajpur and Kendrapara districts, Odisha; purple villages are classified high-risk and blue villages low-risk by the Odisha Flood Hazard Atlas, gray lines are community-development-block boundaries, black lines are district boundaries, and unshaded polygons are non-sampled villages. Panel B plots the flooding rate in each group: bars are the Atlas calibration, the mean across villages of the category-midpoint number of inundated years divided by the 18-year Atlas record (2001–2018); diamonds are survey-reported rates, the share of village-years 2019–2025 in which at least half of the village's surveyed households report economically damaging flooding. Panel C plots village-level balance: each row is the coefficient from a regression of the standardized covariate on the high-risk indicator with block fixed effects, with 95% confidence intervals from heteroskedasticity-robust standard errors. The upper section contains demographic, infrastructure, and wealth measures (from the SHRUG and the survey). The lower block contains physical geography measures (elevation, slope, and drainage from terrain data; soils from SoilGrids; rice suitability from GAEZ).

3.3 Estimation

To measure the effects of living in a high-flood-risk environment, we regress:

$$Y_{ivbt} = \beta \; \textit{High risk}_{vb} + \delta \; \textit{Flood}_{ivbt} + \theta \; (\textit{High risk} \times \textit{Flood})_{ivbt} + \mathbf{X}'_{vb}\gamma + \alpha_{b,t} + \varepsilon_{ivbt} \tag{4}$$

Y_{ivbt} is an outcome for household i in village v located inside block b and surveyed in year t; our main outcome of interest is agricultural net returns.[11] We have three key regressors of interest. $\textit{High risk}_{vb}$ is an indicator for a village designated as high or very high risk in the Flood Hazard Atlas. $\textit{Flood}_{ivbt}$ is an indicator for a damaging flood in village v in year t, defined as at least 50% of survey households *excluding* household i reporting damaging flooding.[12] $\textit{High risk} \times \textit{Flood}_{ivbt}$ is the interaction between the two. $\mathbf{X}_{vb}$ is a vector of village characteristics, including elevation, land slope, drainage, distance to river, rice suitability, and soil characteristics, to improve precision and address any remaining concern that areas with flood risk are worse for agriculture through mechanisms that are independent of propensity to flood. $\alpha_{b,t}$ are block-$\times$-wave fixed effects which non-parametrically absorb any differences across blocks in each survey wave. Finally, ε_{ivbt} is an error term; we cluster at the village level. In this specification, β is identified from comparisons between nearby low- and high-risk villages within the same block and wave, that are similar on key observables but differentiated in flood risk owing to micro-scale topographical variation. δ and θ are identified off of quasi-random flood events in villages within the same risk category, block, and wave.

To interpret β as the causal effect of flood risk on our outcomes of interest, low- and high-risk villages must be otherwise similar. Our localized sampling strategy is designed to achieve this, and Panel C of Figure 1 provides evidence that our villages do not differ on key observables. Moreover, Odisha's restrictions on agricultural land markets strongly limit the main channel through which farmer sorting typically occurs. It is also possible that the same land characteristics that generate flood risk could independently lower net returns; we therefore include detailed measures of eleva-

[11] Because our survey does not capture every possible component of agricultural profits, we use the terminology "net returns" in our empirical exercise, though the theoretical object of interest is profits.

[12] Because the survey question asks farmers about *damaging* floods, we exclude household i's own report to ensure that the event indicator does not mechanically relate to household i's own outcome. The definition of a flood in a village should therefore be seen as inundation that causes positive damages to a substantial fraction of (surveyed) households.

tion, drainage, soils, river proximity, and agricultural suitability in our preferred specification, and present robustness to adjusting the control set.

To interpret δ and $\delta+\theta$ as the effect of a realized flood on returns in low- and high-risk villages, realized flooding must be an exogenous shock, unrelated to other determinants of agricultural returns after conditioning on risk group, block-$\times$-wave fixed effects, and controls. This seems reasonable, given that farmer actions do not cause floods, and advance warning remains limited. From a measurement point of view, however, we might be concerned that farmers who saw a reduction in profits may be more likely to report that a flood occurred. To guard against this, we use a leave-one-out definition of flooding, excluding the farmer's own flood reports from the treatment variable. Finally, flooding may lower farm profits through mechanisms other than crop destruction. Our estimates will be net of these channels and net of adaptation, which is appropriate for our purposes. We present additional robustness checks in Section 4.4.

Link to accounting framework The four village-season cells implied by Equation (4) are low-risk ordinary seasons (the omitted category), high-risk ordinary seasons (β), low-risk flood seasons (δ), and high-risk flood seasons ($\beta+\delta+\theta$). The mapping to our accounting framework is therefore

$$C_N = -\beta, \qquad D_L = -\delta, \qquad D_H = -(\delta+\theta).$$

We call C_N the all-season component because β enters the high-risk outcome in both ordinary and flood seasons; we identify it in ordinary seasons because the realized-flood terms vanish when $F = 0$. Combining these estimates with the Atlas flood frequencies yields the total burden of high flood risk relative to low flood risk, $C = C_N + p_H D_H - p_L D_L$, which we decompose into the all-season component, C_N, and the realized-flood component, $C_F = p_H D_H - p_L D_L$.

4 Results

4.1 Flood risk and agricultural returns

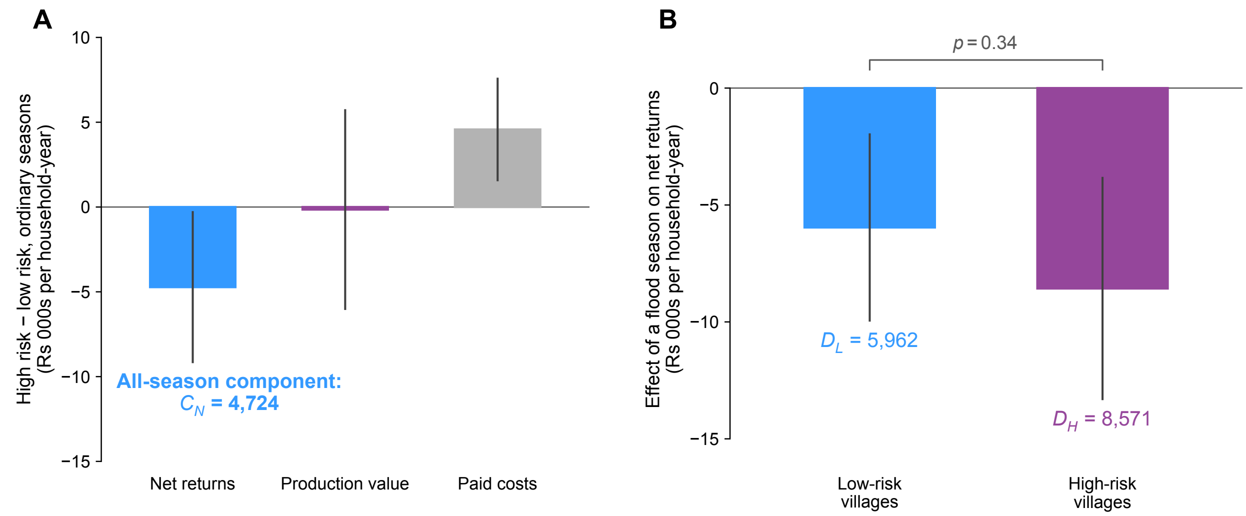
We begin by documenting that living in a high-flood-risk village is associated with lower agricultural returns. Panel A of Figure 2 plots the ordinary-season gap between high- and low-risk villages in net returns (blue), value of production (purple), and paid costs (gray). The negative of the net-returns effect is C_N, the all-season component in our accounting framework. Panel B plots the effect of a flood in low-risk villages (blue) and in high-risk villages (purple). These effects are $-D_L$ and $-D_H$, respectively.

This figure demonstrates that high flood risk affects agriculture in two ways. First, flood risk lowers net returns in all seasons by 4,724 Rs. ($p = 0.040$).[13] These impacts are large: approximately 21% of the ordinary-season low-risk mean returns of 22,044 Rs. This gap is driven by differences in spending on the farm: 4,574 Rs. ($p = 0.004$); we see no all-season effect on the value of production (point estimate: -150 Rs., $p = 0.960$).

Second, floods are highly damaging. In low-risk villages, realized floods reduce net returns by $D_L = 5{,}962$ Rs. ($p = 0.004$), or 27% of the observed low-risk ordinary-season mean. In high-risk villages, floods reduce net returns by $D_H = 8{,}571$ Rs. ($p < 0.001$), 42% of the observed high-risk ordinary-season mean. In both sets of villages, flood events have large impacts on agriculture. The point estimate is larger in high-risk villages, but we cannot reject equal flood damages across the two environments ($p = 0.34$); our theory does not sign the difference.

[13] We present robustness checks in Appendix B.

Figure 2: The two inputs to the accounting: the all-season component and the flood loss



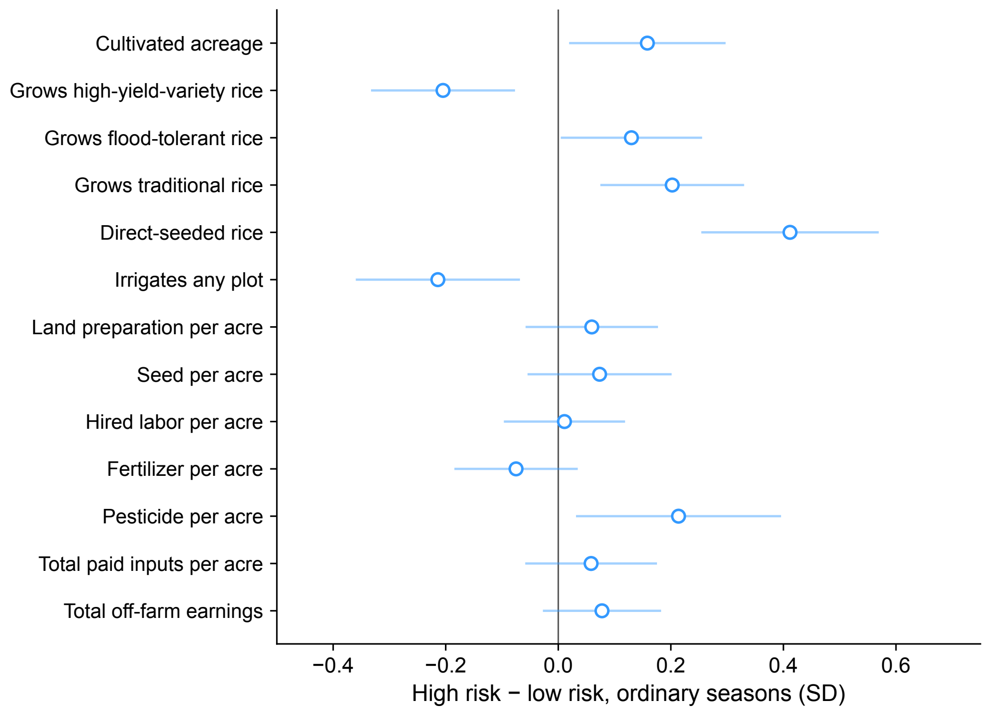
Notes: This figure reports the two estimated inputs to the accounting framework, Equation (3). Panel A plots the high-minus-low contrast in net returns, production value, and paid costs in ordinary seasons (Rs. per household-year; axis in thousands): $\hat{\beta}$ in Equation (4), which controls for block-×-wave fixed effects and village characteristics. The net-returns contrast equals $-C_N$, the (positive) all-season component. Panel B plots the effect of a damaging flood season on net returns in low-risk villages ($\hat{\delta}$) and high-risk villages ($\hat{\delta}+\hat{\theta}$), with the p-value for their equality; the accounting's losses D_L and D_H, stated in the annotations, are the negatives of these effects. A damaging flood season is defined as at least 50% of the other surveyed households in the village reporting economically damaging flooding that season (leave-one-out). Whiskers are 95% confidence intervals from standard errors clustered at the village level.

4.2 Evidence of production reallocation

Flood risk imposes costs in all seasons because farmers respond by adjusting their input choices in order to mitigate losses under inundation, sacrificing returns when no floods occur. Figure 3 provides evidence of this mechanism: high- and low-risk households choose a different bundle of inputs, consistent with defensive adaptation. We estimate effects on inputs using Equation (4) with production choices as the outcome variable; our estimate of interest is $\hat{\beta}$, the gap between high- and low-risk households in ordinary seasons.

We find that households in high-risk villages grow different varieties of rice, the dominant crop, than their low-risk counterparts, planting more flood-tolerant (+0.13 SD, $p = 0.043$) and traditional (+0.20 SD, $p = 0.002$) rice and forgoing high-yield-variety rice (−0.20 SD, $p = 0.002$). Farmers in high-risk areas also engage in more direct rice seeding (+0.41 SD, $p < 0.001$), a practice actively promoted by the government in flood zones. High-risk farmers are also less likely to irrigate their plots (−0.21 SD, $p = 0.004$) and spend somewhat more on pesticide per acre (+0.21 SD, $p = 0.022$), though the gap in their overall expenses per acre is not statistically distinguishable

Figure 3: Production choices differ with flood risk



Notes: This figure plots differences between high- and low-risk villages in production choices and reallocation margins in ordinary seasons: each row is the coefficient on the high-risk indicator (i.e., $\hat{\beta}$) from Equation (4), which controls for block-×-wave fixed effects and village characteristics, scaled by the standard deviation of the outcome among low-risk ordinary-season observations. Whiskers are 95% confidence intervals from village-clustered standard errors.

from zero. High-risk farmers also cultivate more land, partially offsetting the fact that returns per acre are 4,791 Rs. (37%, $p < 0.001$) lower in ordinary seasons than in low-risk villages. We see no offsetting differences in off-farm earnings. Taken together, high-risk farmers appear to reallocate inputs z consistent with the framework in Section 2: they spend 17% more, farming 13% more land at the same input intensity, but the value of their output is similar to their low-risk counterparts. C_N is thus consistent with households reallocating investments towards flood protection.

4.3 The economic burden of flood risk

Finally, Figure 4 and Table 1 carry our main estimates through our accounting framework and present the headline results. Low-risk villages have expected annual returns of 21,867 Rs. Realized flooding reduces net returns by 5,962 Rs. in low-risk villages (D_L) and 8,571 Rs. in high-risk villages (D_H). Floods are rare in low-risk villages (p_L), only occurring in 3% of seasons, and common in high-risk villages (p_H), taking place in 57% of seasons. Thus, the annual expected gap in returns

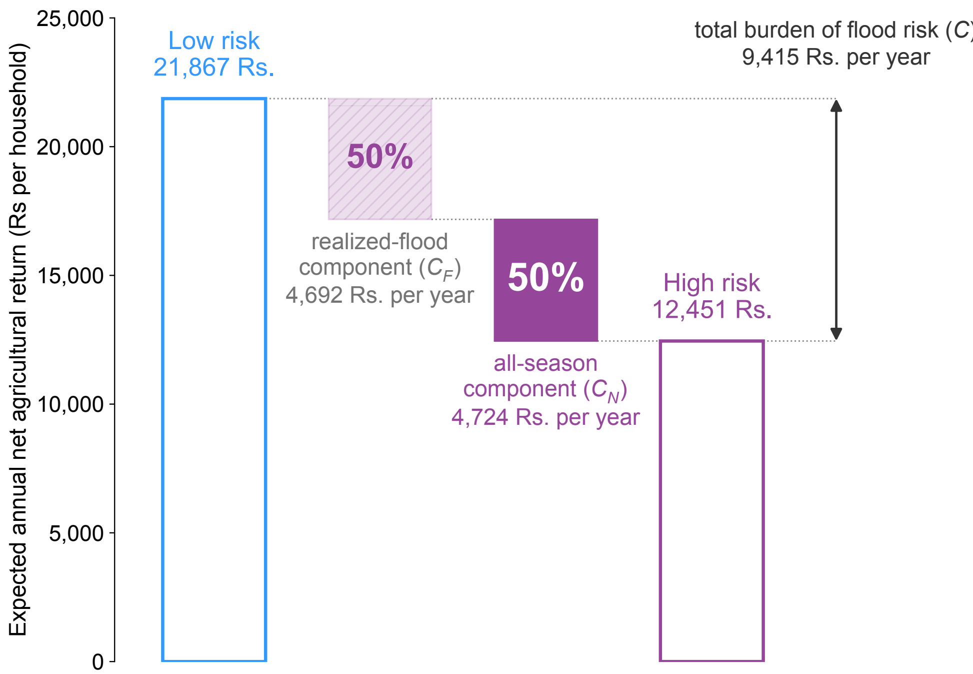
Table 1: Flood risk in our accounting framework

Quantity	Estimate	Uncertainty
Panel A. Inputs to the annual accounting		
All-season component, C_N	4,724	(2,287)
Flood-season loss, low-risk villages, D_L	5,962	(2,054)
Flood-season loss, high-risk villages, D_H	8,571	(2,436)
p-value for $D_L - D_H = 0$: 0.34		
Flood-frequency gap, Δp	0.538	
high-risk rate, p_H	0.568	
low-risk rate, p_L	0.030	
Panel B. Annual burden accounting		
Low-risk expected annual return, $E_L = \mu_L(0) - p_L D_L$	21,867	
All-season component, C_N	*4,724*	(2,287)
Realized-flood component, $C_F = p_H D_H - p_L D_L$	*4,692*	(1,368)
Annual burden, $C = C_N + C_F$	*9,415*	(1,839)

Notes: This table constructs the overall effect of living in a high-risk vs. low-risk village on agricultural net returns. Panel A reports the ingredients. C_N is the all-season profit gap between high- and low-risk environments (Section 2), $-\hat{\beta}$ in Equation (4). D_L and D_H are the reduction in net returns from a realized flood (where our flood treatment is defined as having at least 50% of survey respondents (excluding household i) reporting flooding) in low- and high-risk villages, $-\hat{\delta}$ and $-\hat{\delta} - \hat{\theta}$ from Equation (4). The estimation sample consists of 3,995 household-wave observations: 902 low- and 735 high-risk observations in ordinary seasons, and 709 low- and 1,649 high-risk observations in flood seasons. Δp is the difference in flood frequencies between high- and low-risk villages, measured using the Flood Hazard Atlas. Panel B shows the total agricultural burden of flood risk. We begin with the expected annual return in low-risk villages, $E_L = \mu_L(0) - p_L D_L$, where $\mu_L(0)$ is the low-risk ordinary-season mean. The all-season component is C_N, and the realized-flood component is $C_F = p_H D_H - p_L D_L$. The total expected annual burden is $C = C_N + C_F$. In the Uncertainty column, parentheses contain standard errors clustered at the village level. Units are in Rs. per household-year except for the probabilities.

between high- and low-risk villages, C, is 9,415 Rs., or 43% of low-risk expected annual returns. 4,692 Rs. (50%) of this gap comes from the greater expected burden of realized floods in high-risk villages, the realized-flood component. 4,724 Rs. (50%) comes from the all-season component, C_N. Our preferred calibration of flood frequencies uses the Flood Hazard Atlas, which provides a long history. Instead using flood histories from the survey yields a total burden of 8,377 Rs., of which 56% occurs in ordinary seasons, compared with 9,415 Rs. and 50% under our preferred calibration. These results make clear that (i) as highlighted in prior work, floods are extremely damaging; and (ii) the cost of flood *risk* is substantially larger than the effects of realized floods alone.

Figure 4: The burden of flood risk



Notes: This figure plots the overall burden of living with flood risk. The far-left bar shows expected annual returns per household in the low-risk environment, $E_L = \mu_L(0) - p_L \times D_L$, where $\mu_L(0)$ is average ordinary-season returns in the low-risk environment, p_L is the probability the low-risk environment experiences a flood, and D_L are the damages when a flood occurs. The light purple inner block shows the realized-flood component of the burden, $C_F = p_H D_H - p_L D_L$, and the dark purple inner block shows the all-season component, C_N. We present the total burden, C, in the arrow at right. Percentages are the share of C that is attributable to each component. The far-right bar shows expected annual returns in the high-risk environment.

4.4 Robustness checks

Appendix B presents additional results and tests that directly stress our identifying assumptions. We first assess the robustness of the all-season component, C_N. Appendix Figure B.1 replaces our binary high/low risk classification with more granular underlying Flood Hazard Atlas categories. We find that the ordinary-season gap on net returns and high-yield-variety rice cultivation is largest in the Atlas's highest-risk villages, consistent with a dose-response relationship in flood risk. An Oster (2019) exercise further suggests that unobservables would need to be more than 100 times more important than our controls, and operate in the opposite direction, to explain away the all-season net returns gap (Oster selection ratio = -131). Appendix Figure B.2 shows that the shift in ordinary-season returns extends through the bottom 75% of the distribution, rather than being driven by outliers. To directly address concerns that land quality impacts net returns other than through flood risk, and to address other possible confounders, Appendix Figure B.3 shows that our estimate of C_N is insensitive to the sequential addition of controls for elevation and slope, drainage and rivers, soil quality, and rice suitability, and their functional form. The results are also robust to adding controls for household size, education, and owned land, or choosing controls through double-selection LASSO (Belloni et al., 2014). Appendix Table B.1 shows that our estimate of C_N is stable across alternative definitions of realized flooding and alternative constructions of net returns, is not driven by lagged flood events, and estimating effects on an ordinary-season-only sample if anything increases the estimated gap. Finally, using Conley (1999) standard errors rather than clustering by village (per Conley and Kelly (2025)) does not meaningfully change the estimated uncertainty.

Additionally, Appendix Table B.2 shows that the impacts of a realized flood are similar under alternative definitions of a damaging flood, and robust to a panel specification with household fixed effects, which absorbs all time-invariant household (and thus village) heterogeneity. Appendix Figure B.3 demonstrates that the realized flood impacts are insensitive to controls. To test whether realized floods are orthogonal to village characteristics, we pool across all villages and estimate the effects of flooding on net returns using both our cross-sectional specification and a specification that includes household fixed effects. The two yield nearly identical results: floods reduce net returns by 7,410 Rs. in the cross-sectional specification and 7,550 Rs. in the household fixed effects approach. As in the cross-sectional specification, our estimates of D_H and D_L are noisier. Though their relative magnitude reverses, they remain statistically indistinguishable ($p = 0.61$); both specifications imply that floods cause losses of a similar order of magnitude. As the types of confounders that are netted out by household fixed effects, such as land quality or farmer ability, likely impact *both* flood-season returns *and* ordinary-season returns, these results are also reassuring about the identification of β. Finally, Appendix Table B.4 shows that the headline accounting is insensitive to the calibration of p_H and p_L, including measuring these values using the survey rather than the Atlas.

5 Conclusion

Floods are among the most damaging environmental hazards, and their costs are disproportionately borne by the poor (Hallegatte et al., 2017; Tellman et al., 2021; Food and Agriculture Organization of the United Nations, 2023b). In this paper, we document that the costs of living with flood *risk* are substantially larger than the impacts of realized flooding alone. We design a survey of households in flood-prone Odisha, India, to measure the costs of flood risk on agricultural returns, using a cross-sectional research design enabled by our sample of villages within small geographic areas facing different flood risk due to hyper-local fluctuations in geography. We demonstrate that flood risk meaningfully depresses agricultural returns, caused *both* by frequent, damaging floods *and* by reduced returns in all seasons, including those without floods. Despite being largely left unmeasured, this latter mechanism has roughly the same effect on expected annual returns as the former.

Our findings reveal that reducing flood risk is worth substantially more than the costs of realized floods alone suggest. The setting we study – smallholder farmers on a floodplain – describes millions of people in low- and middle-income countries (Lowder et al., 2021; Rentschler et al., 2022). Climate change will worsen flood risk for many millions of people: the population at risk of river flooding alone is expected to more than double under a 2°C warming scenario, and increase five-fold under a 4°C warming scenario (Caretta et al., 2022); these increases will be concentrated in Africa and Asia (Hirabayashi et al., 2021). Valuing these increases using realized flood damages alone would understate their costs by approximately half.

References

Acampora, Michelle, Lorenzo Casaburi, and Jack Willis (2025): "Land rental markets: Experimental evidence from Kenya," *American Economic Review*, 115 (3), 727–771.

Asher, Sam, Tobias Lunt, Ryu Matsuura, and Paul Novosad (2021): "Development research at high geographic resolution: An analysis of night-lights, firms, and poverty in India using the SHRUG open data platform," *The World Bank Economic Review*, 35 (4), 845–871.

Balboni, Clare, Johannes Boehm, and Mazhar Waseem (2026): "Firm adaptation in production networks: Evidence from extreme weather events in Pakistan," *American Economic Review*, forthcoming.

Bates, Paul D., Matthew D. Wilson, Matthew S. Horritt, David C. Mason, Nick Holden, and Anthony Currie (2006): "Reach scale floodplain inundation dynamics observed using airborne synthetic aperture radar imagery: Data analysis and modelling," *Journal of Hydrology*, 328 (1–2), 306–318.

Belloni, Alexandre, Victor Chernozhukov, and Christian Hansen (2014): "Inference on treatment effects after selection among high-dimensional controls," *The Review of Economic Studies*, 81 (2), 608–650.

Bradt, Jacob T. and Joseph E. Aldy (2025): "Private benefits from public investment in climate adaptation and resilience," NBER Working Paper 33633, National Bureau of Economic Research.

Burke, Marshall and Kyle Emerick (2016): "Adaptation to climate change: Evidence from US agriculture," *American Economic Journal: Economic Policy*, 8 (3), 106–140.

Burlig, Fiona, Amir Jina, Erin Kelley, Gregory Lane, and Harshil Sahai (2024): "The value of forecasts: Experimental evidence from India," NBER Working Paper 32173, National Bureau of Economic Research.

Caretta, Martina A., Aditi Mukherji, Md Arfanuzzaman, Richard A. Betts, Alexander Gelfan, Yukiko Hirabayashi, Tabea K. Lissner, Junguo Liu, Elena Lopez Gunn, Ruth Morgan, Stephen Mwanga, and Seree Supratid (2022): "Chapter 4: Water," in *Climate Change 2022: Impacts, Adaptation and Vulnerability. Contribution of Working Group II to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*, ed. by Hans-Otto Pörtner, Debra C. Roberts, Melinda Tignor, Elvira S. Poloczanska, Kristie Mintenbeck, Andrés Alegría, Marlies Craig, Stephanie Langsdorf, Sophie Löschke, Vera Möller, Andrew Okem, and Bidhan Rama, Cambridge, UK and New York, NY, USA: Cambridge University Press, 551–712.

Carleton, Tamma, Amir Jina, Michael Delgado, Michael Greenstone, Trevor Houser, Solomon Hsiang, Andrew Hultgren, Robert E. Kopp, Kelly E. McCusker, Ishan Nath, James Rising, Ashwin Rode, Hee Kwon Seo, Arvid Viaene, Jiacan Yuan, and Alice Tianbo Zhang (2022): "Valuing the global mortality consequences of climate change accounting for adaptation costs and benefits," *The Quarterly Journal of Economics*, 137 (4), 2037–2105.

Caunedo, Julieta and Namrata Kala (2021): "Mechanizing agriculture," NBER Working Paper 29061, National Bureau of Economic Research.

Choi, Jungyill and Judson W. Harvey (2014): "Relative significance of microtopography and vegetation as controls on surface water flow on a low-gradient floodplain," *Wetlands*, 34 (1), 101–115.

Christian, Paul, Felipe Dunsch, Jonas Heirman, Erin Kelley, Florence Kondylis, Gregory Lane, Jennifer Waidler, Nidhila Adusumalli, Odbayar Batmunkh, and Kriti Malhotra (2025): "Fast action for floods: RCT evidence on forecast-based cash transfers from Bangladesh and Nepal," NBER Working Paper 34414, National Bureau of Economic Research.

Clarke, Daniel J. and Stefan Dercon (2016): *Dull disasters? How planning ahead will make a difference*, Oxford: Oxford University Press.

Conley, Timothy G. (1999): "GMM estimation with cross sectional dependence," *Journal of Econometrics*, 92 (1), 1–45.

Conley, Timothy G. and Morgan Kelly (2025): "The standard errors of persistence," *Journal of International Economics*, 153, 104027.

Debnath, Sridip, Ashok Mishra, Damodhara R. Mailapalli, and Narendra S. Raghuwanshi (2021): "Assessment of rice yield gap under a changing climate in India," *Journal of Water and Climate Change*, 12 (4), 1245–1267.

Deininger, Klaus, Songqing Jin, and Meilin Ma (2022): "Structural transformation of the agricultural sector in low- and middle-income economies," *Annual Review of Resource Economics*, 14, 221–241.

Deininger, Klaus, Songqing Jin, and Hari K. Nagarajan (2009): "Determinants and consequences of land sales market participation: Panel evidence from India," *World Development*, 37 (2), 410–421.

del Ninno, Carlo, Paul A. Dorosh, and Lisa C. Smith (2003): "Public policy, markets and household coping strategies in Bangladesh: Avoiding a food security crisis following the 1998 floods," *World Development*, 31 (7), 1221–1238.

Deryugina, Tatyana and Solomon Hsiang (2017): "The marginal product of climate," NBER Working Paper 24072, National Bureau of Economic Research.

Donovan, Kevin (2021): "The equilibrium impact of agricultural risk on intermediate inputs and aggregate productivity," *The Review of Economic Studies*, 88 (5), 2275–2307.

Emerick, Kyle, Alain de Janvry, Elisabeth Sadoulet, and Manzoor H. Dar (2016): "Technological innovations, downside risk, and the modernization of agriculture," *American Economic Review*, 106 (6), 1537–1561.

European Space Agency (2022): "Copernicus DEM – Global Digital Elevation Model (GLO-30)," ESA Copernicus Data Space Ecosystem, 30 m resolution; produced from WorldDEM (© DLR 2010–2014, © Airbus Defence and Space 2014–2018).

Fernando, A. Nilesh (2022): "Shackled to the soil? Inherited land, birth order, and labor mobility," *Journal of Human Resources*, 57 (2), 491–524.

Fischer, Günther, Freddy O. Nachtergaele, Harrij T. van Velthuizen, Federica Chiozza, Guido Franceschini, Marc Henry, Douglas Muchoney, and Sylvia Tramberend (2021): "Global Agro-Ecological Zones v4: Model documentation," Tech. rep., Food and Agriculture Organization of the United Nations, Rome.

Fisher, Anthony C., W. Michael Hanemann, Michael J. Roberts, and Wolfram Schlenker (2012): "The economic impacts of climate change: Evidence from agricultural output and random fluctuations in weather: Comment," *American Economic Review*, 102 (7), 3749–3760.

Food and Agriculture Organization of the United Nations (2023a): "FAOSTAT statistical database: Crops and livestock products," https://www.fao.org/faostat.

——— (2023b): *The impact of disasters on agriculture and food security 2023: Avoiding and reducing losses through investment in resilience*, Rome: FAO.

Government of India, Ministry of Agriculture and Farmers Welfare (2015): "Compensation for crop damage due to natural calamities," Press Information Bureau, Government of India.

——— (2023): "Annual Report 2022–23," Directorate of Economics and Statistics, New Delhi.

Government of Odisha (1960): "The Orissa Land Reforms Act, 1960 (Odisha Act 16 of 1960)," As amended; Inspector General of Registration, Odisha.

Hallegatte, Stephane, Adrien Vogt-Schilb, Mook Bangalore, and Julie Rozenberg (2017): *Unbreakable: Building the resilience of the poor in the face of natural disasters*, Climate Change and Development Series, Washington, DC: World Bank.

Hirabayashi, Yukiko, Masahiro Tanoue, Orie Sasaki, Xudong Zhou, and Dai Yamazaki (2021): "Global exposure to flooding from the new CMIP6 climate model projections," *Scientific Reports*, 11, 3740.

Hornbeck, Richard (2012): "The enduring impact of the American Dust Bowl: Short- and long-run adjustments to environmental catastrophe," *American Economic Review*, 102 (4), 1477–1507.

Hornbeck, Richard and Suresh Naidu (2014): "When the levee breaks: Black migration and economic development in the American South," *American Economic Review*, 104 (3), 963–990.

Hsiang, Solomon (2016): "Climate econometrics," *Annual Review of Resource Economics*, 8 (1), 43–75.

Hultgren, Andrew, Tamma Carleton, Michael Delgado, Diana R Gergel, Michael Greenstone, Trevor Houser, Solomon Hsiang, Amir Jina, Robert E Kopp, Steven B Malevich, et al. (2025): "Impacts of climate change on global agriculture accounting for adaptation," *Nature*, 642 (8068), 644–652.

Innes, Robert (2003): "Crop insurance in a political economy: An alternative perspective on agricultural policy," *American Journal of Agricultural Economics*, 85 (2), 318–335.

IPCC (2022): *Climate Change 2022: Impacts, adaptation and vulnerability. Contribution of Working Group II to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*, Cambridge, UK and New York, NY: Cambridge University Press.

Jagnani, Maulik and Rohini Pande (2024): "Forecasting fate: Experimental evaluation of a flood early warning system," Working paper.

Jia, Ruixue, Xiao Ma, and Victoria Wenxin Xie (2026): "Expecting floods: Firm entry, employment, and aggregate implications," *American Economic Journal: Macroeconomics*, 18 (3), 42–76.

Karlan, Dean, Robert Osei, Isaac Osei-Akoto, and Christopher Udry (2014): "Agricultural decisions after relaxing credit and risk constraints," *The Quarterly Journal of Economics*, 129 (2), 597–652.

Kocornik-Mina, Adriana, Thomas K. J. McDermott, Guy Michaels, and Ferdinand Rauch (2020): "Flooded cities," *American Economic Journal: Applied Economics*, 12 (2), 35–66.

Kousky, Carolyn (2019): "The role of natural disaster insurance in recovery and risk reduction," *Annual Review of Resource Economics*, 11, 399–418.

Lane, Gregory (2024): "Adapting to climate risk with guaranteed credit: Evidence from Bangladesh," *Econometrica*, 92 (2), 355–386.

Lindsay, John B. (2016): "Efficient hybrid breaching-filling sink removal methods for flow path enforcement in digital elevation models," *Hydrological Processes*, 30 (6), 846–857.

Lowder, Sarah K., Marco V. Sánchez, and Raffaele Bertini (2021): "Which farms feed the world and has farmland become more concentrated?" *World Development*, 142, 105455.

Mearns, Robin and Saurabh Sinha (1999): "Social exclusion and land administration in Orissa, India," Policy Research Working Paper 2124, World Bank, Washington, DC.

Mendelsohn, Robert, William Nordhaus, and Daigee Shaw (1994): "The impact of global warming on agriculture: A Ricardian analysis," *American Economic Review*, 84 (4), 753–771.

Merz, Bruno, Günter Blöschl, Sergiy Vorogushyn, Francesco Dottori, Jeroen C. J. H. Aerts, Paul Bates, Miriam Bertola, Matthias Kemter, Heidi Kreibich, Upmanu Lall, and Elena Macdonald (2021): "Causes, impacts and patterns of disastrous river floods," *Nature Reviews Earth & Environment*, 2 (9), 592–609.

Merz, Bruno, Heidi Kreibich, Reimund Schwarze, and Annegret Thieken (2010): "Review article: Assessment of economic flood damage," *Natural Hazards and Earth System Sciences*, 10, 1697–1724.

Mobarak, Ahmed Mushfiq and Mark R. Rosenzweig (2013): "Informal risk sharing, index insurance, and risk taking in developing countries," *American Economic Review*, 103 (3), 375–380.

National Remote Sensing Centre (2019): "Flood Hazard Atlas for Odisha State: A geospatial approach," Tech. rep., National Remote Sensing Centre, Indian Space Research Organisation, Hyderabad, India, prepared in association with the Odisha State Disaster Management Authority (OSDMA), Government of Odisha, and in coordination with the National Disaster Management Authority (NDMA), Government of India.

NITI Aayog (2016): "Report of the Expert Committee on Land Leasing," Tech. rep., Government of India, New Delhi, chair: T. Haque.

OpenStreetMap contributors (2024): "OpenStreetMap data," https://www.openstreetmap.org, river network retrieved via the Overpass API; © OpenStreetMap contributors, ODbL.

Oster, Emily (2019): "Unobservable selection and coefficient stability: Theory and evidence," *Journal of Business & Economic Statistics*, 37 (2), 187–204.

Patel, Dev (2024): "Floods," Working paper, Harvard University.

Poggio, Laura, Luis M. de Sousa, Niels H. Batjes, Gerard B. M. Heuvelink, Bas Kempen, Eloi Ribeiro, and David Rossiter (2021): "SoilGrids 2.0: Producing soil information for the globe with quantified spatial uncertainty," *SOIL*, 7 (1), 217–240.

Rentschler, Jun, Melda Salhab, and Bramka Arga Jafino (2022): "Flood exposure and poverty in 188 countries," *Nature Communications*, 13 (1), 3527.

Rosenzweig, Mark R. and Hans P. Binswanger (1993): "Wealth, weather risk and the composition and profitability of agricultural investments," *Economic Journal*, 103 (416), 56–78.

Rosenzweig, Mark R. and Kenneth I. Wolpin (1985): "Specific experience, household structure, and intergenerational transfers: Farm family land and labor arrangements in developing countries," *The Quarterly Journal of Economics*, 100 (Supplement), 961–987.

Schlenker, Wolfram and Michael J. Roberts (2009): "Nonlinear temperature effects indicate severe damages to U.S. crop yields under climate change," *Proceedings of the National Academy of Sciences*, 106 (37), 15594–15598.

Scown, Murray W., Martin C. Thoms, and Nathan R. De Jager (2015): "Measuring floodplain spatial patterns using continuous surface metrics at multiple scales," *Geomorphology*, 245, 87–101.

Tellman, Beth, Jonathan A. Sullivan, Catherine Kuhn, Albert J. Kettner, Colin S. Doyle, G. Robert Brakenridge, Tyler A. Erickson, and Daniel A. Slayback (2021): "Satellite imaging reveals increased proportion of population exposed to floods," *Nature*, 596, 80–86.

van Ittersum, Martin K., Kenneth G. Cassman, Patricio Grassini, Joost Wolf, Pablo Tittonell, and Zvi Hochman (2013): "Yield gap analysis with local to global relevance – a review," *Field Crops Research*, 143, 4–17.

Ward, Philip J., Brenden Jongman, Frederiek Sperna Weiland, Arno Bouwman, Rens van Beek, Marc F. P. Bierkens, Willem Ligtvoet, and Hessel C. Winsemius (2013): "Assessing flood risk at the global scale: Model setup, results, and sensitivity," *Environmental Research Letters*, 8 (4), 044019.

The burden of flood risk

Online appendix

Fiona Burlig, Jack N. Glaser, and Anant Sudarshan

Contents

A Data

Appendix Table A.1 reports summary statistics for the analysis sample, pooling the 2023 and 2025 survey waves. Our data cover the kharif (summer monsoon) planting season, which coincides with the majority of flooding in Odisha. Panel A provides demographic information. Our households have 6 people on average, and the average maximum educational attainment for a household member is 10.7 years. These households are relatively poor, with an average asset value of 27,128 Rs. (329 USD), and current savings of 19,088 Rs. (231 USD).

Panel B presents summary statistics about flooding in our sample. 60% of our villages are classified as high-risk. Per the Flood Hazard Atlas (National Remote Sensing Centre, 2019), low-risk villages flood only 3% of years, while high-risk villages flood 57% of years over a nearly 20-year history. The Atlas measures flooding via remote sensing, combining data from the "Indian Remote Sensing Satellites and microwave satellite data from Radarsat, RISAT-1 & Sentinel-1 series," in conjunction with river gauge data to match peak flooding. In order to classify a village as having flooded, India's National Remote Sensing Centre compares satellite data on water to standing water coverage. Our survey respondents report more damaging flooding within our sampled years: a damaging flood event occurred for 44% of low-risk and 69% of high-risk household-wave observations pooled across 2023 and 2025.

Panel C summarizes our aggregate revenue and cost estimates, and provides details on land under cultivation. We collected detailed data on each household's three largest plots, which cover 96% of total cultivated acres. On average, the sample is made up of small farmers with a mean cultivated acreage (in the top three plots) of approximately 2 acres. We define our main outcome variable – net agricultural returns – as the total value of kharif production less paid costs. To calculate production value, we gathered data on crop type and output at the plot level. To convert physical output measures to monetary terms we multiply by the median sale price reported by all farmers in a block. Thus valuation is at common prices regardless of whether or how much the household sold.[14] In our preferred specification, we value output at crop-specific prices.

On the cost side, we aggregate over survey-reported expenditures on seed, fertilizer (urea, DAP, potash, and other), pesticides, hired labor, machinery, irrigation, transport and marketing, and a miscellaneous expenses category. We measure pesticides and irrigation at the plot level, seed and transport at the crop level, and fertilizer, labor, machinery, and miscellaneous expenses at the household level. We aggregate all expenditures to the household level. In our main estimates, we include machinery expenditure in 2023 only, because we did not field this survey module in 2025. We also exclude family labor, as it is not directly priced. Panel B of Appendix Table B.1 presents robustness of our main all-season return estimate under alternative constructions of the net returns variable. The point estimates remain relatively stable as we vary these choices.

Finally, Panel D provides information on several input choice variables. Just over half of farmers cultivate traditional rice varieties, with 47% cultivating high-yield varieties and just 5% planting flood-tolerant rice, though with high variance. This may be because, while effective at protecting against floods, it typically sells for lower output prices. A significant minority use direct seeding methods to plant. Odisha is characterized by rain-fed agriculture, so perhaps unsurprisingly, only

[14]To limit a small number of extreme outliers, we winsorize production value and costs at the 1st and 99th percentile within each survey wave, and top-code yields at three times published estimates for each crop (Government of India, Ministry of Agriculture and Farmers Welfare, 2023; Food and Agriculture Organization of the United Nations, 2023a). Agronomists put potential yields in Odisha at about 1.5 times realized yields (Debnath et al., 2021; van Ittersum et al., 2013), so allowing twice that bound is conservative. Winsorizing at the 5th/95th percentiles, or dropping outliers rather than winsorizing, does not change the results. We also estimate prices at the block level, not block-by-wave level, in our preferred specification, as this allows us to construct prices over more observations. Instead using block-by-wave median prices does not change the results.

30% of the sample reports irrigating any plot.

Table A.1: Summary statistics

Variable	Mean	SD	Low-risk mean	High-risk mean	N
Panel A. Household characteristics (household-wave)					
Household size	5.89	2.96	5.98	5.83	3,967
Maximum household education (years)	10.69	3.66	10.57	10.77	3,948
Total asset value (Rs.)	27,128	26,832	27,969	26,560	4,002
Current savings (Rs.)	19,088	33,174	20,235	18,319	3,949
Panel B. Flood exposure and realized flooding					
High-risk village (village)	0.60	0.49	0.00	1.00	401
Atlas annual flood frequency (village)	0.35	0.28	0.03	0.57	401
Damaging-flood event, q=.50 (household-wave)	0.59	0.49	0.44	0.69	3,995
Panel C. Agricultural outcomes (household-wave)					
Net agricultural returns (Rs.)	15,188	35,219	19,472	12,299	4,002
Production value (Rs.)	43,112	41,216	44,115	42,435	4,002
Paid production costs (Rs.)	27,924	24,557	24,643	30,137	4,002
Cultivated acreage (top-3 plots, acres)	2.07	1.56	1.87	2.21	4,002
Net returns per cultivated acre (Rs.)	7,913	19,565	10,874	5,922	3,960
Panel D. Production choices (household-wave)					
Grows traditional rice	0.51	0.50	0.47	0.54	4,002
Grows high-yield-variety rice	0.47	0.50	0.52	0.43	4,002
Grows flood-tolerant rice	0.05	0.21	0.03	0.06	4,002
Direct-seeded rice (majority of rice area)	0.31	0.46	0.19	0.40	3,800
Irrigates any plot	0.30	0.46	0.36	0.26	3,966
Fertilizer expenditure per acre (Rs.)	2,907	2,466	3,113	2,769	3,960
Paid inputs per acre (Rs.)	15,905	15,366	15,684	16,053	3,960

Notes: This table reports summary statistics for the June–September kharif (summer monsoon) season for the analysis sample: 401 villages and 4,002 household-wave observations with full geographic covariates, pooling the 2023 and 2025 survey waves (attrition between samples is about 13%). Low- and high-risk are defined per the Flood Hazard Atlas definition we describe in the main text. Panels A, C, and D are household-wave measures; in Panel B, the high-risk indicator and the Atlas annual flood frequency (Flood Hazard Atlas category midpoints over 2001–2018) are village-level constants, and the damaging-flood event is the household-wave indicator used in estimation. A damaging flood season is defined for each household as one in which at least 50% of the other sampled households in its village report crop flood damage. N varies with item availability. Units are Rs. per household-year for monetary variables unless otherwise indicated.

B Additional results and robustness checks

B.1 Effects by intensity of flood risk

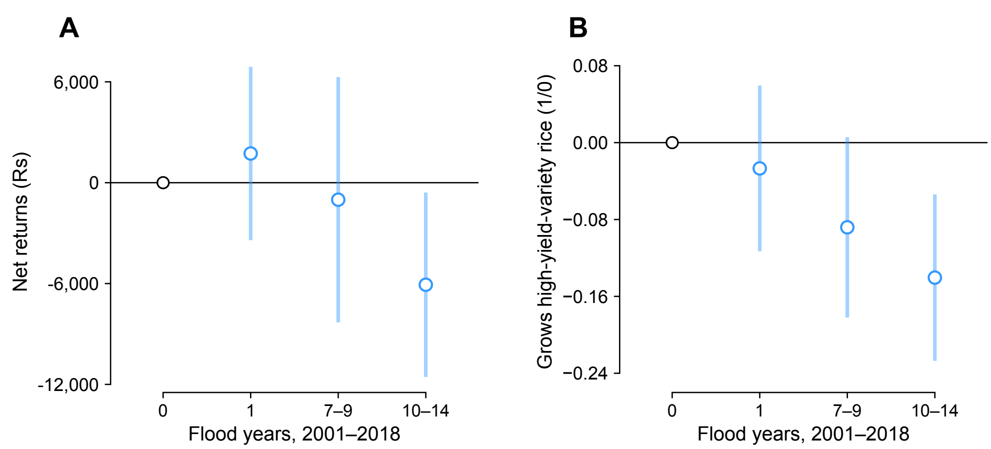
The Flood Hazard Atlas that we use to define low and high flood risk categorizes villages into bins based on the number of years that they experienced flooding between 2001 and 2018: no flooding (which do not appear in the Atlas), 1 flood year, 2–4 flood years, 5–6 flood years, 7–9 flood years, and 10–14 flood years. The Atlas uses objective measures – satellite-based inundation and river gauge data – to determine flooding. In our main estimates, we combine the no flooding and 1 flood year groups into our "low risk" category, and combine the 7–9 and 10–14 flood year groups into our "high risk" category. To maximize power given our budget, we did not sample villages from the 2–4 or 5–6 flood year groups.

Appendix Figure B.1 presents the effect of being in each Flood Hazard Atlas risk group on net returns (Panel A) and growing high-yield-variety rice (Panel B). To estimate these impacts, we use a modified version of Equation (4):

$$Y_{ivbt} = \sum_{g \in \mathcal{G}} \Big[\beta^g \, \mathbf{1}[Risk\ group\ = g]_{vb} + \theta^g \, (\mathbf{1}[Risk\ group\ = g] \times Flood)_{ivbt} \Big] + \delta \ Flood_{ivbt} + \mathbf{X}'_{vb}\gamma + \alpha_{b,t} + \varepsilon_{ivbt} \tag{B.1}$$

where $g \in \mathcal{G}$ are the finer risk groups described above, and all other terms are identical to those in Equation (4). No flooding is the omitted category, so all effects are relative to this group. The gap in net returns is driven by the highest-risk group; the probability of planting high-yield-variety rice declines monotonically as risk increases.

Figure B.1: Net returns and high-yield-variety rice planting across Flood Hazard Atlas risk categories

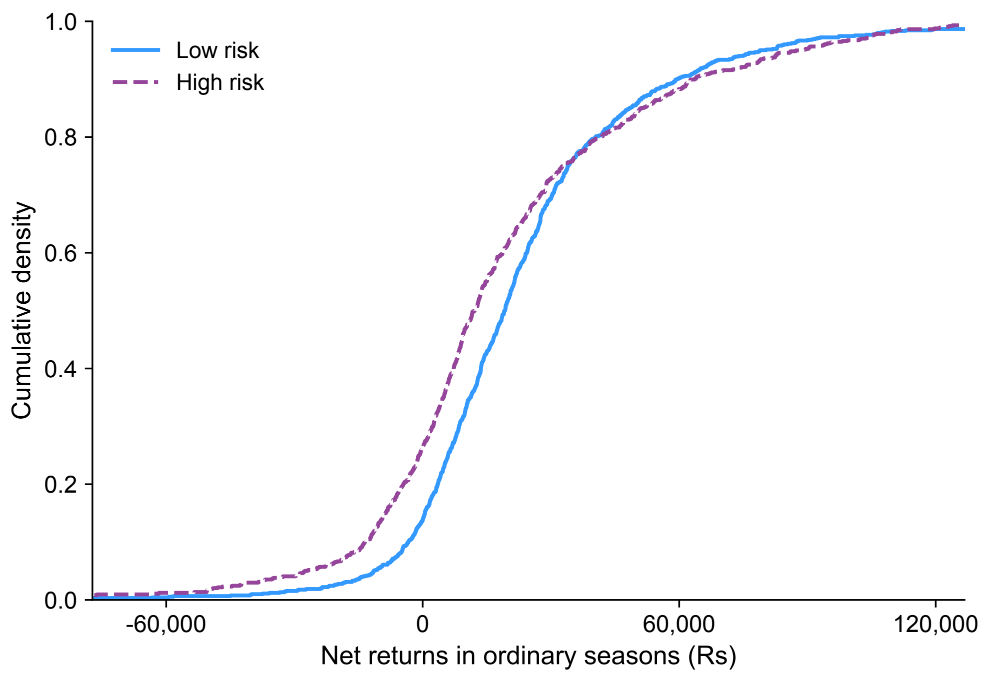


Notes: This figure plots the effect of flood risk on net agricultural returns (Panel A) and the probability of cultivating high-yield-variety rice (Panel B), measured using the finest-available risk groups from the Flood Hazard Atlas. There are 75 no-flooding villages, 85 villages with 1 flood year from 2001 to 2018, 108 villages with 7–9 flood years, and 133 villages with 10–14 flood years. Our sample deliberately excluded villages with 2–4 or 5–6 flood years. All estimates come from a version of Equation (4) with the binary high-risk indicator replaced by group indicators, which includes block-×-wave fixed effects and the core terrain and soil controls. No-flood villages are the omitted category (black circle). Whiskers are 95% confidence intervals from standard errors clustered at the village level.

B.2 Distribution of ordinary-season net returns effects

To evaluate where in the distribution we see effects of flood risk on ordinary-season net returns, Appendix Figure B.2 plots the CDF of net returns separately for low- (solid blue) and high-risk (dashed purple) villages, using a sample of ordinary seasons only. We residualize out block-×-wave fixed effects and the core set of terrain and soil control variables, adding the grand mean back in for display purposes. We observe a gap between the two risk categories through approximately the 75th percentile of the distribution, suggesting that the effect of risk is a broad shift of the distribution, and not driven by outliers. The impact of risk on net returns is if anything largest at the bottom of the distribution, consistent with the most flood-exposed households within high-risk villages making the largest input-bundle adjustments.

Figure B.2: The effect of flood risk on ordinary-season net returns: Distribution



Notes: This figure plots empirical CDFs of net agricultural returns for low-risk (solid blue) and high-risk (dashed purple) villages in ordinary seasons, after residualizing block-×-wave fixed effects and the core terrain and soil controls. We truncate the horizontal axis at the 0.5th and 99th percentiles for display purposes.

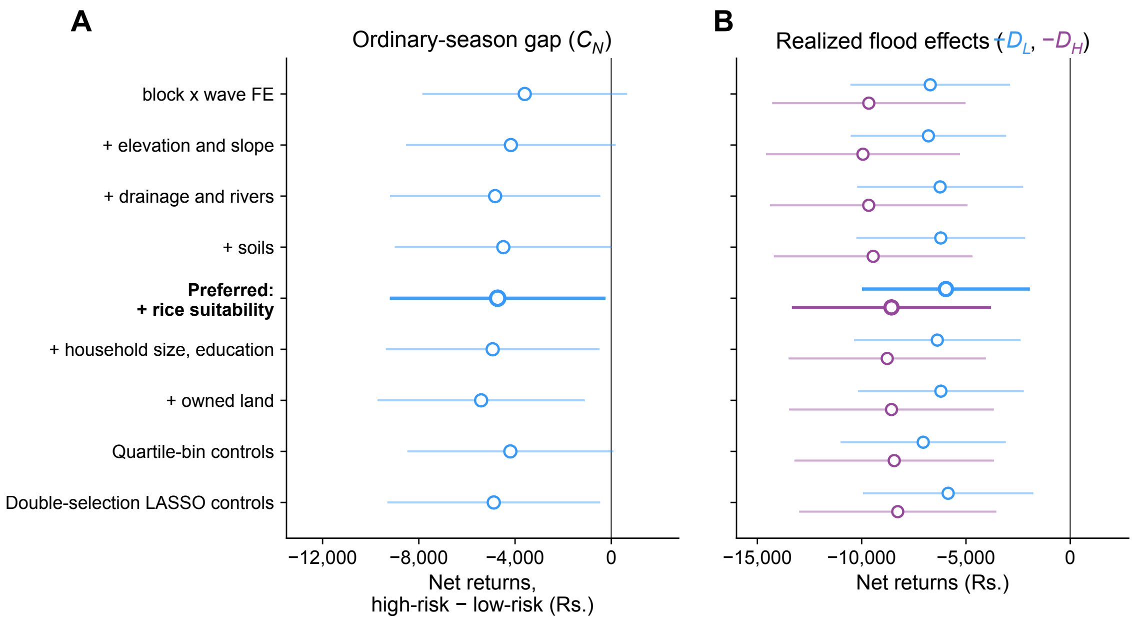
B.3 Robustness to choice of controls

Appendix Figure B.3 presents how our estimates of the all-season component of the net returns gap (C_N, Panel A) and the impacts of flooding in low- and high-risk villages ($-D_L$, $-D_H$, Panel B) vary with our chosen set of control variables. All estimates include block-×-wave fixed effects, which we show alone in the top row. The next four rows progressively add geographic/geological controls: elevation and slope, drainage and rivers, soils, and rice suitability (our preferred specification). Then, we add household characteristics: first size and education, then the amount of land owned. In the penultimate row, we vary functional form, including quartile bins of each of the controls included in the preferred specification. The last row selects controls using double-selection LASSO (Belloni et al., 2014), where the candidates include the controls in the preferred specification, additional geographic controls, demographics from the 2011 Census, and 2009–2013 DMSP nightlights. The

LASSO selects six: the share of village land less than 2 meters above the nearest minor and major rivers, distance to minor and major river, distance to town, and rice suitability. Across all specifications, point estimates for all three objects are relatively stable, assuaging concerns that our results are driven by a particular set of covariates or highly vulnerable to omitted variables.

Following Oster (2019), adding controls moves the ordinary-season estimate away from zero (from 5,689 Rs. to 6,739 Rs. estimated on ordinary-season observations only, using block-×-wave fixed effects throughout), so selection on unobservables proportional to selection on observables implies a larger gap: the bias-adjusted estimate at $R^2_{max} = 1.3\tilde{R}^2$ is 7,491 Rs., and driving the estimate to zero would require Oster's $\delta = -131$ – unobservables more than one hundred times as important as the controls, selecting in the opposite direction.

Figure B.3: Robustness: Alternative controls



Notes: This figure reports the two estimated inputs to the accounting under progressively richer control sets. Panel A plots the ordinary-season net-returns gap between high- and low-risk villages ($-C_N$); Panel B plots the effect of a flood season on net returns, separately in low-risk ($-D_L$, blue) and high-risk ($-D_H$, purple) villages, both estimated using Equation (4). Every rung includes block-×-wave fixed effects. The second, third, fourth, and fifth rows add controls cumulatively – elevation and slope; drainage and rivers; soils; and rice suitability – until the preferred specification (bold), which is the specification in Table 1. Next, we add household size and education, and subsequently also the amount of land owned by each household, to the preferred specification. In the penultimate row, we instead replace every control in the preferred specification with within-sample quartile-bin indicators. Finally, we select controls using a double-selection LASSO procedure (Belloni et al., 2014). Markers are point estimates and whiskers are 95% confidence intervals from standard errors clustered at the village level. Across rungs, C_N ranges from 3,598 to 5,410 Rs., D_L from 5,855 to 7,049 Rs., and D_H from 8,271 to 9,938 Rs.

B.4 Robustness of the all-season component (C_N)

Appendix Table B.1 presents additional robustness checks on the all-season component of the net returns gap, C_N. In all but the final row, which we describe below, we estimate treatment effects using Equation (4).

To interpret β as the all-season return gap, it must isolate the impact of high vs. low flood risk on net returns during ordinary seasons (see Equation 1). Panel A therefore varies our definition of a flood season. In our primary specification, we set the flood indicator equal to 1 if 50% of households (other than i) report damaging flooding. As expected, making this more (less) stringent by requiring

75% (25%) of households to report flooding increases (decreases) the size of the coefficient.

Panel B reports how our estimate changes as we adjust the definition of the outcome variable. In the first row after the primary estimate, we estimate net returns per cultivated acre. In the second, third, and fourth rows, we change the construction of revenue: using only sales to value output, using variety-(rather than crop-)specific output prices, or using block-×-wave-median rather than block-median output prices. The estimate falls when only valuing sales, suggesting that production for home consumption is important to the difference between high- and low-risk villages, but is otherwise stable. Next, in the fifth, sixth, seventh, eighth, and ninth rows, we vary how we define costs: valuing family labor at 80% of the hired wage (Deininger et al., 2022; Caunedo and Kala, 2021), valuing land costs, excluding machinery costs from both waves, excluding transport and marketing costs, or excluding miscellaneous other expenses. Excluding machinery reduces the all-season component of the gap, implying that machinery spending (e.g., tractor rental) is an important difference between low- and high-risk villages, but the point estimates are otherwise stable across choices around costs. This is corroborated by the 2023-only row in Panel D.

Panel C presents standard errors using the Conley (1999) spatial HAC; this does not meaningfully impact inference.

Finally, Panel D presents three sample restrictions. First, we limit to ordinary seasons only; this is equivalent to fully interacting all controls in Equation (4) with the flood indicator. The point estimate rises substantially, suggesting that the imposed homogeneity in the effects of the controls is conservative. Second, we keep only villages that reported no flood in the two years prior to the survey wave. Again, the point estimate rises, assuaging concerns that C_N is a lagged effect of flooding rather than an equilibrium outcome. Finally, we restrict the sample to the 2023 wave only, and replace the block-×-wave fixed effects in Equation (4) with block fixed effects only. The point estimate is meaningfully larger in 2023. This is consistent with us having only captured machinery spending in this earlier wave.

Table B.1: Robustness of the all-season returns component (C_N)

Specification	Estimate	(SE)
Panel A. Alternative damaging-flood-event definitions		
Primary: q=.50 leave-one-out event	*-4,724*	*(2,287)*
q=.75 event	-4,848	(2,065)
q=.25 event	-3,210	(2,531)
Continuous leave-one-out share (gap at share 0)	-3,794	(2,354)
Panel B. Outcome construction (q=.50 event)		
Primary: net agricultural returns	*-4,724*	*(2,287)*
Per cultivated acre	-4,791	(1,170)
Sales-based output valuation	-3,183	(1,617)
Variety-specific output prices	-4,942	(2,320)
Block-×-wave median output prices	-4,902	(2,287)
Family labor valued at 80% of hired wage	-4,218	(2,463)
Land costs included (rent + sharecrop comp.)	-3,961	(2,153)
Machinery costs excluded (both waves)	-3,379	(2,595)
Transport and marketing costs excluded	-4,693	(2,308)
Miscellaneous other expenses excluded	-4,812	(2,293)
Panel C. Spatial correlation		
Primary estimate, spatial (Conley) SEs, 10 km	-4,724	(2,377)
Panel D. Sample restriction		
Ordinary seasons only (flood-season obs. dropped)	-6,739	(2,372)
No damaging flood in the two preceding years	-6,182	(2,725)
2023 wave only (block fixed effects)	-6,692	(2,441)

Notes: This table reports the coefficient on the high-risk indicator – the ordinary-season high–low contrast, so that C_N is the negative of the net-returns estimate – under alternative definitions of the damaging-flood event (Panel A), alternative constructions of the return outcome (Panel B), adjusted uncertainty (Panel C), and restricted samples (Panel D). Every row (except the first row of Panel D) includes the realized flooding indicator, its interaction with the high-risk indicator, the core terrain and soil controls, and block-×-wave fixed effects, with standard errors clustered at the village level. In Panel A, the primary specification defines a flood as 50% of *other* survey respondents in the village-wave reporting flooding; the q=.75 and q=.25 rows raise or lower the threshold share of other sampled households reporting crop flood damage; and the continuous row replaces the binary event with the leave-one-out damage share, reporting the high-risk contrast at share zero. In Panel B, the primary outcome is net returns as defined in the main text; alternatives divide by the household's current cultivated acreage, use sales-based output valuation, and change the set of costs that are included in the net returns calculation. In the primary specification, we have 401 villages and 3,995 household-wave observations; N varies slightly with outcome availability, and Panel D's restricted samples are smaller. The variety-priced row revalues rice output at block-median prices computed separately by seed-variety class (from growers of a single class), rather than a single paddy price. The block-×-wave-priced row values output at block-median prices computed separately by survey wave, rather than pooled across waves. Panel C reports the primary estimate with standard errors allowing arbitrary spatial correlation among villages within 10 km (Conley, 1999; Conley and Kelly, 2025); 2 km and 5 km cutoffs (not shown) give similar results. Panel D restricts the sample. In the first row, we estimate the gap using ordinary-season observations only (equivalent to fully interacting the regression specification with the flood indicator). The second row restricts to villages with no damaging flood in the two years preceding the survey season, so that the ordinary-season contrast cannot reflect recovery from recent floods. The third row restricts to the 2023 year alone, where we observe machinery spending.

B.5 Robustness of the realized flood effects (D_L, D_H)

Appendix Table B.2 presents robustness of the effects of a realized flood on net returns in low- and high-risk villages, $-D_L$ and $-D_H$. In Panel A, we vary the definition of a realized flood. Throughout Panel A, we estimate the low- and high-risk-specific effects using Equation (4), and estimate the pooled effect using a pooled version:

$$Y_{ivbt} = \beta \ \mathit{High\ risk}_{vb} + \delta \ \mathit{Flood}_{ivbt} + \mathbf{X}'_{vb}\gamma + \alpha_{b,t} + \varepsilon_{ivbt} \tag{B.2}$$

where all terms are defined as in Equation (4), but we do not estimate separate effects of idiosyncratic flooding by risk category.

In our primary specification, we set the flood indicator equal to 1 if 50% of households (other than i) report damaging flooding. Making this more or less stringent does not substantially change the conclusion that realized flooding is damaging in both low- and high-risk villages.

In Panel B, we instead estimate the impacts of idiosyncratic flooding using a specification with household fixed effects, to assuage concerns that flooding is systematically correlated with ability, productivity, or other omitted variables:

$$Y_{ivbt} = \delta \ \mathit{Flood}_{ivbt} + \theta \ (\mathit{High\ risk} \times \mathit{Flood})_{ivbt} + \eta_i + \alpha_{b,t} + \varepsilon_{ivbt} \tag{B.3}$$

where η_i is a household fixed effect and all other terms are as in Equation (4); we also estimate an equivalent version of Equation (B.2) to estimate pooled impacts. The pooled effect is very similar to the primary specification. While the low- and high-risk-village-specific effects switch order relative to our primary specification, they remain statistically indistinguishable, and the conclusion that flooding is damaging is preserved.

Table B.2: Robustness of the effect of a realized flood

Event specification	Estimate	(SE)
Panel A. Cross-sectional specification		
Primary (q=.50 event)		
Pooled, all villages	-7,410	(1,840)
Low-risk villages (D_L)	*-5,962*	*(2,054)*
High-risk villages (D_H)	*-8,571*	*(2,436)*
p-value, low = high	[0.34]	
q=.75 event		
Pooled, all villages	-8,255	(1,673)
Low-risk villages	-6,591	(2,128)
High-risk villages	-9,297	(2,147)
p-value, low = high	[0.33]	
q=.25 event		
Pooled, all villages	-7,804	(1,956)
Low-risk villages	-5,580	(2,021)
High-risk villages	-10,173	(2,726)
p-value, low = high	[0.10]	
Continuous leave-one-out share (0 to 1)		
Pooled, all villages	-10,758	(2,187)
Low-risk villages	-8,744	(2,424)
High-risk villages	-12,289	(2,818)
p-value, low = high	[0.25]	
Panel B. Household fixed effects (balanced panel)		
q=.50 event		
Pooled, all villages	-7,550	(2,223)
Low-risk villages	-8,264	(2,467)
High-risk villages	-6,773	(2,858)
p-value, low = high	[0.61]	

Notes: This table reports the effect of a damaging flood season on net agricultural returns. In Panel A, we estimate treatment effects using Equation (4), which includes block-×-wave fixed effects and the core terrain and soil controls. In Panel B, we instead estimate treatment effects from a specification which includes household fixed effects and block-×-wave fixed effects. In both panels, standard errors (in parentheses) are clustered at the village level. Coefficients are the raw flood-season effects, so a negative net-returns entry is a loss; the accounting's D_L and D_H are the negatives of the italicized primary net-returns entries. For each event definition, effects are reported separately in low- and high-risk villages from the interacted regression, with the p-value for their equality below. Pooled rows report the flood coefficient from Equation (B.2). The q=.75 and q=.25 rows change the threshold share of other sampled households required to report a flood in order for a household to be considered treated. The continuous row replaces the binary event with the leave-one-out reported-damage share and is the effect of moving that share from 0 to 1 – a full-village damage report – not damage per flood event. In Panel B, we use the primary q=.50 definition throughout. We estimate treatment effects on a balanced panel of 3,708 household-waves within 371 villages. 1,854 households enter the household fixed-effects specification: 7 households lack the leave-one-out event indicator in one wave; 23 villages, surveyed only in 2023, exit the panel entirely.

B.6 The effect of flood risk and realized flooding on migration

We estimate the impact of flood risk and realized flooding on migration using Equation (4); we present the results in Appendix Table B.3. Columns (1)–(3) refer to migration in the current kharif season. Seasonal migration is non-trivial among low-risk villages during ordinary seasons, with 21% of households sending at least one migrant out of the household. However, neither flood risk nor flooding (either in low- or high-risk villages) has an economically meaningful or statistically significant impact on the probability of sending migrants out of the home, the amount of time spent away, or remittances. Columns (4) and (5) refer to more permanent migration in the past 10 years, which is exceedingly rare: only 0.3% of low-risk ordinary-season observations report doing this. While risk does not impact longer-term migration, we find evidence that households report sending children away to protect them from flooding when realized flooding occurs.

Table B.3: Migration responses to flood risk and realized flooding

	(1) Any member moved away	(2) Person-days away	(3) Money sent home (Rs.)	(4) Migrated member out of village	(5) Sent child away
High-risk gap, ordinary seasons	-0.016	-2.5	237	0.006	0.004
	(0.023)	(2.8)	(550)	(0.005)	(0.004)
Flood effect, low-risk villages	-0.021	-3.6	408	0.009	0.017
	(0.024)	(2.6)	(636)	(0.007)	(0.008)
Flood effect, high-risk villages	0.002	0.2	12	0.004	0.022
	(0.022)	(2.7)	(528)	(0.005)	(0.005)
Low-risk ordinary-season mean	0.206	21.0	3,461	0.003	0.003
Observations	3,958	3,958	3,958	3,992	3,995

Notes: This table reports the effect of flood risk and of realized flooding on migration, estimated using Equation (4). Columns (1)–(3) refer to the current kharif season; Columns (4) and (5) refer to the migration that has occurred in response to flooding any time in the last 10 years. The first row reports the coefficient on the high-risk indicator, the second row reports the coefficient on the flood indicator, and the third row reports the sum of the coefficient on the flood indicator with the coefficient on the interaction between the flood indicator and the high-risk indicator. All regressions include block-×-wave fixed effects and the core terrain and soil controls, with standard errors (in parentheses) clustered at the village level.

B.7 Alternative calibrations of p_L and p_H

In our primary accounting, we measure p_L and p_H, the probability of flooding in low- and high-risk environments, using the Flood Hazard Atlas. This is our preferred approach because the Atlas analyzes a long history of flooding, spanning nearly two decades, which allows us to best approximate the long-run risk of inundation. Appendix Table B.4 presents robustness of the realized-flood component of the net returns gap (C_F), the total gap (C), and the share of the gap coming from the all-season component (C_N/C) to alternative calibrations of p_L and p_H. First, we use flood histories as reported by our survey, which cover 2019–2025. The estimates of p_L and p_H are both meaningfully higher than their Atlas equivalents (0.367 vs. 0.030 and 0.682 vs. 0.568, respectively), but the overall conclusions are similar: the realized-flood component and the overall net returns gap remain large, and the share of the net returns gap attributable to the all-season component is if anything slightly higher. Next, we show three adjustments using the Atlas itself. The Atlas assigns each village to a risk group based on the number of flood years it experienced between 2001 and 2018, but rather than reporting an exact number of years of inundation, it reports a range. We therefore assign all villages to the lower (higher) end of the range in the third (fourth) row.

Finally, the Atlas does not include villages with no flooding. Using village identifiers from the Census (Asher et al., 2021), we match all villages to the Atlas. We label villages that do not match as having zero floods from 2001 to 2018. The last row excludes these from our calibration. None of the three Atlas adjustments meaningfully changes our conclusions: floods are damaging, flood risk is damaging, and the all-season share of the burden is roughly half.

Table B.4: The burden accounting under alternative flood-probability calibrations

Calibration	p_L	p_H	C_F	C	C_N/C
Primary: Atlas category midpoints	*0.030*	*0.568*	*4,692*	*9,415*	*0.50*
Survey-reported flood histories, 2019–2025	0.367	0.682	3,654	8,377	0.56
Atlas category lower bounds	0.030	0.482	3,950	8,674	0.54
Atlas category upper bounds	0.030	0.655	5,433	10,156	0.47
Atlas midpoints, excluding never-flooded villages from p_L	0.056	0.568	4,538	9,262	0.51

Notes: This table recomputes the accounting under alternative calibrations of the flood probabilities p_L and p_H. The estimated components C_N, D_L, and D_H are held fixed across rows; only the calibration changes, with $C_F = p_H D_H - p_L D_L$ and $C = C_N + C_F$. The rows are therefore alternative mappings from the estimated components into the total burden, not competing specifications: inference for C in the main specification is conditional on the chosen calibration, and this table is the sensitivity analysis for uncertainty about that choice. Our preferred calibration uses Flood Hazard Atlas category midpoints, which are independently measured and based on a longer historical record (2001–2018). The survey-history calibration is especially informative because it is more recent and because its definition of flooding – households' reports of economically damaging floods, with a village-year coded as flooded when at least half of surveyed households report damage – matches the concept used to estimate D_L and D_H, whereas the Atlas measures inundation. The final row computes p_L using only low-risk villages that appear in the Atlas ($p_L = 1/18$), rather than coding villages absent from the Atlas as never flooded. Across the calibrations shown, the total burden ranges from 8,377 to 10,156 Rs. and the all-season share from 47% to 56%.