Basis Risk and Low Demand for Weather Index Insurance

Kodama, Wataru; Miura, Ken; Sakurai, Takeshi

KYOTO UNIVERSITY

NRE Discussion Papers No. 2021-02
November 2021
Division of Natural Resource Economics, Graduate School of Agriculture, Kyoto University

Basis Risk and Low Demand for Weather Index Insurance

Wataru Kodama, Ken Miura*, and Takeshi Sakurai

November 10, 2021

Faculty/Graduate School of AGRICULTURE KYOTO UNIVERSITY

Division of Natural Resource Economics
Kitashirakawa Oiwake-cho, Sakyo-ku, Kyoto 606-8502, Japan

Basis risk and low demand for weather index insurance*

Wataru Kodama† Ken Miura‡ Takeshi Sakurai§

November 2021

Abstract

Basis risk — an imperfect correlation between an aggregate index and idiosyncratic crop damage — has been widely recognized as a major impediment to embracing index insurance. However, empirical evidence is still scarce because of the difficulty in its direct measurement. In this study, we estimate the impact of the basis risk on demand for a rainfall index insurance product using household survey data from rural Zambia. First, we develop a simple model of insurance demand to motivate our econometric specifications. Then, we quantify the basis risk for each surveyed household with past rainfall data at the plot-level. Exploiting changes in insurance design across years, we use within-household variations in the basis risk to estimate its impact. Empirical results illustrate that the basis risk has significant and adverse effects on insurance demand. Despite its statistical significance, our results also suggest that minimizing the basis risk would not yield enough economic benefits to offset the associated costs.

Keywords: Weather index insurance, basis risk, rational demand model, Zambia

JEL Classification: D81, O12, O16, Q14.

*We thank Hisaki Kono, Kazushi Takahashi, Chieko Umetsu, and the seminar participants at Kyoto University and the Society for Environmental Economics and Policy Studies (SEEPS) Annual Conference 2021 for their valuable comments. Errors, if any, are our own.

†Kyoto University. E-mail: kodama.wataru.87a@st.kyoto-u.ac.jp

‡Kyoto University. Corresponding author: Division of Natural Resource Economics, Graduate School of Agriculture, Kyoto University, Kitashirakawa-oiwakecho, Sakyo-ward, Kyoto city, Kyoto 606-8502, Japan; E-mail: miura.ken.3e@kyoto-u.ac.jp

§University of Tokyo. E-mail: atsakura@mail.ecc.u-tokyo.ac.jp

1 Introduction

A decade has passed since Giné et al. (2008) and Cole et al. (2013) reported a low uptake rate of the "innovative" weather index insurance product in India and discussed potential barriers to spurring the potential demand.[1] However, reported demand for index insurance in developing countries remains low despite potentially large welfare benefits (see for e.g., Cai, 2016; Hill et al., 2019; Janzen and Carter, 2018). A recent review from J-PAL (2016) reports that the uptake rate of unsubsidized insurance products ranges from 6% to 18%. Although many studies attribute the major constraint to basis risk — the risk of no insurance payout when farmers incur loss because of an imperfect correlation between the index and actual output loss — empirical evidence is still scarce because of difficulties in measuring the basis risk (e.g., Clarke, 2016; Hazell and Hess, 2010; Jensen et al., 2016; Janzen et al., 2020).[2] Using insurance sales data from a drought-prone region in Zambia, in this study, we provide a direct test of the impact of the basis risk on farmers' actual insurance purchases.

Low uptake and cover rates are matters of concern for both policymakers and insurance providers.[3] From a policy perspective, the extent to which the expected loss would be covered by the insurance product is a critical question, because it affects households' expected utility in the short run and investment and consumption levels eventually. If farmers are inelastic to a change in the basis risk, policy support for its reduction (e.g., installing more weather stations) would not be cost-effective. On the supply side, selling an index insurance product in developing countries is a small-profit-and-quick-return type of business. Thus, scaling up insurance demand is necessary to sustain the

[1]Farm households in developing countries are subject to substantial weather risks, which causes their production and consumption levels to fluctuate. Such risks may also prevent farmers from profitable investments (Cai, 2016; Hill et al., 2013; Janzen and Carter, 2018; Karlan et al., 2014). Therefore, an agricultural insurance policy to mitigate these risks is a promising policy tool for farmers in developing countries. In the index insurance scheme, payouts are conditional only on officially observable indexes such as rainfall. This feature minimizes information asymmetry problems, as well as the significant transaction costs from which indemnity-based crop insurance suffers (Arnott and Stiglitz, 1991; Hess and Hazell, 2009). Offering protection against weather risks while maintaining lower prices, index insurance has been attracting considerable attention from both researchers and policymakers. See Ali et al. (2020) for a recent review on index insurance in developing countries.

[2]Other suggested barriers include a lack of trust in the insurance provider (Cole et al., 2013; Giné and Yang, 2009), low financial literacy (Cai et al., 2020; Cai and Song, 2017; Gaurav et al., 2011), and liquidity constraints (Cai et al., 2020; Cole et al., 2013; Giné et al., 2008).

[3]Throughout the paper, we define "cover rate" as the coverage from the insurance payouts divided by the expected loss because of weather disasters.

business.

Basis risk has attracted significant attention in the theoretical and empirical literature (e.g., Clarke, 2016; Janzen et al., 2020; Jensen et al., 2016; Yu et al., 2019) not only because it constrains insurance demands but also because it is a unique risk specific to index insurance products. As described by Jensen et al. (2018), basis risk is the “Achilles heel of index insurance”: conditional payments based only on observable indexes enable index insurance products to be affordable but spawn basis risk. Empirical studies typically relate low insurance demand to either spatial or product basis risks (see for e.g., Conradt et al., 2015; Hill et al., 2013; Janzen et al., 2020; Jensen et al., 2016; Mobarak and Rosenzweig, 2012). The former risk arises from the poor correlation of rainfall between individual plots and the weather station, whereas the latter is due to weak yield–weather index correlations. Irrespective of type, the basis risk is widely recognized as a factor that severely suppresses index insurance demand (Clarke, 2016; Hill et al., 2013). Despite significant attention, most empirical studies only provide indirect evidence on the effects of the basis risk because of the difficulties in its direct measurement. For example, some studies use the distance from a farmer's plot to the reference weather station as a proxy of spatial basis risk to estimate its impact on the insurance demand (e.g., Hill et al., 2013; Mobarak and Rosenzweig, 2012). Such empirical proxies may be correlated with unobserved factors that determine insurance demand, which raises concerns about the credibility of the evidence.

To fill this research gap, in this study, we use plot-level rainfall data to provide direct empirical evidence on the impact of the basis risk on actual weather index insurance purchases. During the 2011/12–2013/14 crop seasons, we introduced rainfall index insurance contracts to local farmers in Southern Province, Zambia. Farm households in the study site do not have access to irrigation facilities and are exposed to substantial drought risks. As a unique feature of this study, we have access to plot-level daily rainfall data for a subset of survey households for five consecutive crop seasons between 2007/08 and 2011/12. These rainfall records allow us to estimate a direct measure of spatial basis risk, defined as the probability of false–negative (i.e., no payout when the policyholder incurs loss). Another feature of this research context is that we could exploit within-household variations in the basis risk for estimation because the payout condition specified in the insurance contract has changed over the years. This contributes to controlling for unobserved household characteristics that may determine both the basis risk and insurance demand.

First, we develop a simple model of insurance demand based on Clarke's (2016) framework to motivate our econometric specifications. Using plot-level rainfall data as well as the survey data, we then investigate how rural farmers react to the change in the estimated probability of false–negative as a direct measurement of the basis risk. We also test other model predictions to confirm the validity of our theoretical model and empirical specifications in our context.

Besides rich policy implications, the evidence we provide contributes to a better understanding of the role of the basis risk in insurance demand. First, this is one of the few studies to estimate the household-level basis risk (unlike a distance proxy) and demonstrates direct evidence of the basis risk effects on actual insurance purchases. Jensen et al. (2016, 2018) estimated household-level basis risk. Using panel data from index-based livestock insurance (IBLI) sales in Kenya, they define the "basis error" (the deviation of actual loss from payout) in previous years as an estimate of the basis risk. However, one concern about their approach would be that the basis error may be endogenous to the demand for insurance contracts in previous years (Janzen et al., 2020). Accordingly, we introduce more exogenous and theory-based estimates following Clarke (2016) by defining the empirical counterpart of the basis risk as the probability of false–negative. Assuming that rainfall is a random draw from specific probability distributions, we quantify the basis risk for each surveyed household from the unique plot-level rainfall data and historical rainfall data from the reference weather stations. With this probability measure, we relate the basis risk to insurance units purchased by the household.

Second, this study provides a unique contribution by investigating how households' purchase behaviour responds to changes in the design of insurance contracts, and hence basis risk (cf, Janzen et al., 2020). If households are elastic to the basis risk, policy-makers may encourage insurance uptake and achieve higher cover rates by reducing the basis risk. Thus, confirming whether households react to such an exogenous contract change is more policy-relevant than merely illustrating the cross-sectional correlation between basis risk measures and insurance demand. Changing index thresholds and referred weather stations, we sold different insurance products across the survey years in our insurance sales. Leveraging this exogenous change in the insurance contract design, we could exploit within-household variations in estimated basis risk to evaluate how households respond to basis risk. The results indicate that households respond to the changes in basis risk; on average, a 5 percentage points (p.p.) reduction in the false–negative probability would have led to a 32–38% increase in the unit of purchase.

Third, this study translates theoretical predictions into an empirical framework and tests model validity using data. Although some empirical studies incorporate the role of basis risk in insurance demand into the empirical specification (e.g., Jensen et al., 2016; Janzen et al., 2020; Mobarak and Rosenzweig, 2012), this is one of the few studies specifying econometric models fully based on an optimal behaviour characterized by a rational demand model of weather index insurance. The results indicate that a set of plot-level probabilities of drought and false–negative explains the empirical pattern of insurance demand, consistent with the model predictions.

Finally, our household fixed effect estimation allows us to measure the economic impact of the basis risk in terms of the cover rate. Our results suggest that a 5 p.p. decrease in basis risk leads to a 1.66–2.77 p.p. increase in the coverage of expected loss with the insurance contract. Although the basis risk suitably explains the low demand for weather index insurance, reducing such risks cannot be a cost-effective policy instrument given its modest economic impacts and high costs associated with minimizing the basis risk.

The remainder of this study is structured as follows. Section 2 introduces a simple model of the index insurance demand. Section 3 describes the study's context, survey design, and data. Section 4 presents the econometric specifications to test the model predictions about the basis risk effects and estimation framework of key probabilities for our subsequent empirical analysis. Section 5 provides empirical results and discusses the implications for policy interventions to reduce the basis risk. Section 6 presents our conclusions.

2 Model

In this section, we present a simple model to describe how farmers decide to adopt index insurance contracts in the presence of the basis risk and weather risks.

2.1 Model of Demand for Index Insurance

The economic agent in the proposed model is a farmer who faces a risk of crop loss because of drought. The farmer is strictly risk-averse and their preference is presented by a von Neumann–Morgenstern utility function U satisfying $U' > 0$ and $U'' < 0$. This utility function is defined over their consumption level $X = W + Y$, where Y is the crop output and W denotes wealth endowments. We assume that W is constant ($W = w$), whereas Y is a random variable that takes a value of $\underline{y}$ in a drought year with probability $p \in (0, 1)$ and y in a normal year with probability $1 - p$. Note that y is strictly greater than $\underline{y}$. Given this, the expected utility is:

$$EU(x) = pU(w + \underline{y}) + (1 - p)U(w + y).$$

Now we introduce a weather index insurance product that judges the state as "drought" when the index takes the value I with the probability $q \in (0, 1)$. This probability (q) is not necessarily equal to p because of differences in rainfall patterns between the farmer's plot and reference weather station. Without loss of generality, we further assume that the product compensates a value of 1 for each unit of contract in the case of "drought." Hence, insurance premium per unit is $P = mq$, where $m > 0$ is a multiple. The insurance is actuarially fair when $m = 1$, actuarially unfair when $m > 1$, and actuarially favorable when $m < 1$. If this product has no risk of contractual nonperformance (i.e., a perfect match between the index and farmer's crop loss), then the expected utility when the farmer purchases α units is:

$$EU(x) = pU(w + \underline{y} + \alpha(1 - mq)) + (1 - p)U(w + y - \alpha mq).$$

The farmer chooses their insurance units $\alpha \geq 0$ to maximize their expected utility.

To introduce the concept of basis risk to this setup, let $r \in (0, 1)$ be a probability of false–negative, that is, the case when the farmer incurs a loss but receives no insurance payout. Following Clarke (2016), we specify the probability structure of weather index insurance product defined by

Table 1: Probability structure of weather index insurance

	Index $= 0$	Index $= I$	Total
Output $= \underline{y}$	r	$p - r$	p
Output $= y$	$1 - q - r$	$q + r - p$	$1 - p$
Total	$1 - q$	q	1

Table 2: Pay-off structure of weather index insurance

State	$(\underline{y}, I)$	$(\underline{y}, 0)$	$(y, 0)$	(y, I)
Probability	$p - r$	r	$1 - q - r$	$q + r - p$
Pay-off, no insurance	$w + \underline{y}$	$w + \underline{y}$	$w + y$	$w + y$
Pay-off, purchased α units	$w + \underline{y} + \alpha(1 - mq)$	$w + \underline{y} - \alpha mq$	$w + y - \alpha mq$	$w + y + \alpha(1 - mq)$

(p, q, r) as in Table 1. This set of probabilities holds the following relationship (Clarke, 2016):

$$p - q < r < p(1 - q).$$

This assumption is necessary for all states to have positive probabilities of occurrence. In addition, $r < p(1-q)$ is necessary for the index to well predict the bad harvest, that is, $\frac{\mathrm{P}(\underline{y},I)}{\mathrm{P}(y,I)} > \frac{\mathrm{P}(\underline{y},0)}{\mathrm{P}(y,0)}$. As the probability of false–negative is r, the probability of insurance payment when the farmer incurred loss is discounted by r and is $p - r$. Similarly, the probability of no payment in a normal year is $1 - q - r$. Finally, the probability of false–positive is the difference between q and p with the addition of r. The farmer fully acknowledges this probability structure when making a decision, and thus, this set of probabilities (p, q, r) characterizes their insurance purchase.

Given the probability structure, the farmer potentially faces four different states: output loss with insurance payout $(\underline{y}, I)$; no loss without insurance payout $(y, 0)$; loss without insurance payout $(\underline{y}, 0)$; and no loss with insurance payout (y, I). Table 2 summarizes the pay-off by the state of the world. Therefore, when the farmer purchases $\alpha(\geq 0)$ units of the index insurance contract, the expected utility is:

$$\begin{aligned} EU(x) &= (p-r)U(w+\underline{y}+\alpha(1-mq)) + rU(w+\underline{y}-\alpha mq) \\ &\quad + (1-q-r)U(w+y-\alpha mq) + (q+r-p)U(w+y+\alpha(1-mq)). \end{aligned} \tag{1}$$

The optimal insurance demand can be derived from the first derivative of the above expected utility with respect to α. However, we cannot obtain any closed-form solution even under common preference forms including log utility and constant relative and absolute risk-aversion (CRRA and CARA) utility. Thus, we employ a second-order Taylor approximation around $\bar{x} = E(x)$:

$$\begin{aligned} EU(x) &\approx E[U(\bar{x}) + U'(\bar{x})(x-\bar{x}) + \frac{1}{2}U''(\bar{x})(x-\bar{x})^2] \\ &= U(\bar{x}) + \frac{1}{2}U''(\bar{x})V(x). \end{aligned}$$

Under the concavity assumption of the utility function ($U'' < 0$), this Taylor approximation implies that the farmer faces the trade-off between decreases in the expected consumption and its variance when purchasing an extra unit of insurance. From the pay-off structure presented in Table 2, the expected value and variance of pay-off x are:

$$\begin{aligned} &E(x) = w + p\underline{y} + (1-p)y + (1-m)q\alpha, \\ \text{and } \; &V(x) = q(1-q)\alpha^2 - 2\{p(1-q)-r\}(y-\underline{y})\alpha + p(1-p)(y-\underline{y})^2. \end{aligned}$$

When the insurance product is actuarially fair (i.e., $m = 1$)[4], the expected pay-off becomes $E(x) = w + p\underline{y} + (1-p)y$, independent of purchased units of the insurance contract α. Therefore, the farmer's optimal behavior is to minimize the variance by purchasing the following unit:

$$\alpha^* = \frac{p(1-q)-r}{q(1-q)}(y-\underline{y}) \tag{2}$$

Note that the non-negativity constraint does not bind under the assumption of $r < p(1-q)$. This approximated solution leaves some implications on the basis risk and its impact on the insurance demand. First, the solution implies the linear and negative effect of basis risk. This justifies our

[4]As we argue later, this is the case in our empirical setting.

linear specification in the subsequent empirical framework. Second, the magnitude of the basis risk effect is proportional to the expected loss, that is, the difference between the expected outputs in normal and drought years. Rephrasing it, the cover rate has a linear relationship with the basis risk. As another implication, other probability components (p, q) also affect the insurance demand. Similar to the basis risk, the probability of crop loss p has a linear relationship with the insurance demand, and the impact is proportional to the expected loss. The probability of insurance payout q has a theoretically indeterminate effect on the demand (See Section A.1 in the Appendix for the comparative statics). In our data, p and r are heterogeneous across the surveyed households, whereas q is constant because households purchased the same insurance product in a given year. In addition, the same household faces the different basis risk r across survey years because of the changes in the payout condition of the insurance contract, whereas p is constant. Finally, the optimal demand is independent of risk preference parameters as well as initial endowments. Under the above approximation in which the farmer cares only up to the variance of their pay-off, risk attitudes do not affect the optimal demand. However, this does not hold when the farmer considers higher moments of their pay-off in decision-making.

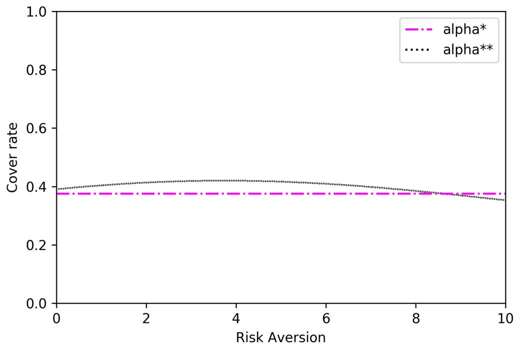
2.2 Numerical solutions

The optimal unit of the insurance contract in equation (2) is an approximated solution. In a general setting, the optimal unit should satisfy the following first-order condition when the insurance is actuarially fair ($m = 1$) based on equation (1):

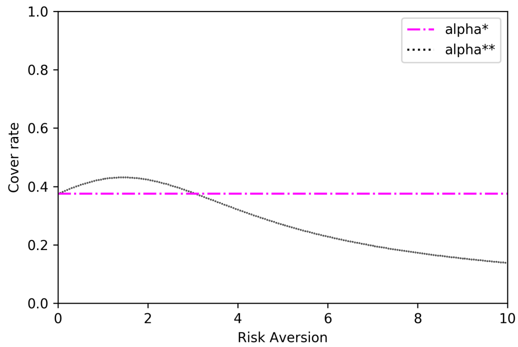
$$\begin{aligned} EU'(x) &= (p-r)(1-q)U'(w+\underline{y}+\alpha(1-q)) - rqU'(w+\underline{y}-\alpha q) \\ &\quad - (1-q-r)qU'(w+y-\alpha q) + (q+r-p)(1-q)U'(w+y+\alpha(1-q)) = 0. \end{aligned}$$

To check the validity of the approximation, we compare the Taylor-approximated solution α^* with a series of numerical solutions, denoted by α^{**}, which satisfy the above condition. In particular, we evaluate the relationship of insurance demand with (i) the degree of risk-aversion and (ii) initial endowments under the CRRA or CARA utility function.

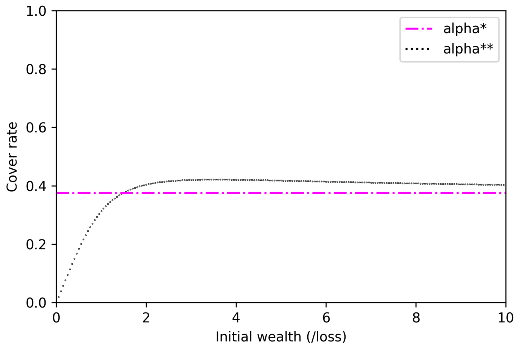
Figure 1 illustrates a pair of approximated and numerical solutions (α^*, α^{**}) by utility func-



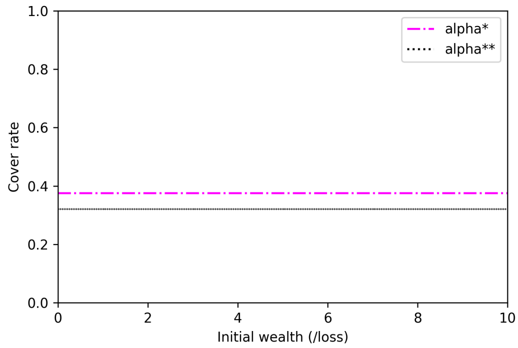
(a) Risk-aversion vs. Insurance demand (CRRA)



(b) Risk-aversion vs. Insurance demand (CARA)



(c) Initial wealth vs. Insurance demand (CRRA)



(d) Initial wealth vs. Insurance demand (CARA)

Figure 1: Approximated and numerical solutions of insurance demand under CRRA and CARA

Notes: The cover rate shows the proportion of insurance payout to loss, $\alpha/(y-\underline{y})$. The specified functional forms are $U(x) = x^{1-\gamma}/(1-\gamma)$ (CRRA) and $U(x) = -\exp(-\gamma x)/\gamma$ (CARA), where γ is a risk preference parameter. To compute numerical solutions, we set $(p, q, r) = (0.2, 0.2, 0.1)$. An initial wealth is assumed to be $w = 3(y-\underline{y})$ for panels (a) and (b) and risk preference parameter is assumed to be $\gamma = 4$ for panels (c) and (d).

tional form. Panels 1a and 1c indicate that α^* provides a good approximation for the numerical solutions under the CRRA utility. By contrast, we observe a discrepancy between approximated and numerical solutions for risk-averse farmers under the CARA utility. The hump-shaped relationship between the risk attitude and insurance demand in Panel 1b is consistent with the Clarke's (2016) general theorem. In the empirical analysis, we test whether the degree of risk-aversion affects the units of actual insurance uptake.

[5]The dry season (May–October) has no rain at all. As most local farmers do not have access to irrigation, agricultural activities are limited to small-scale vegetable cultivation.

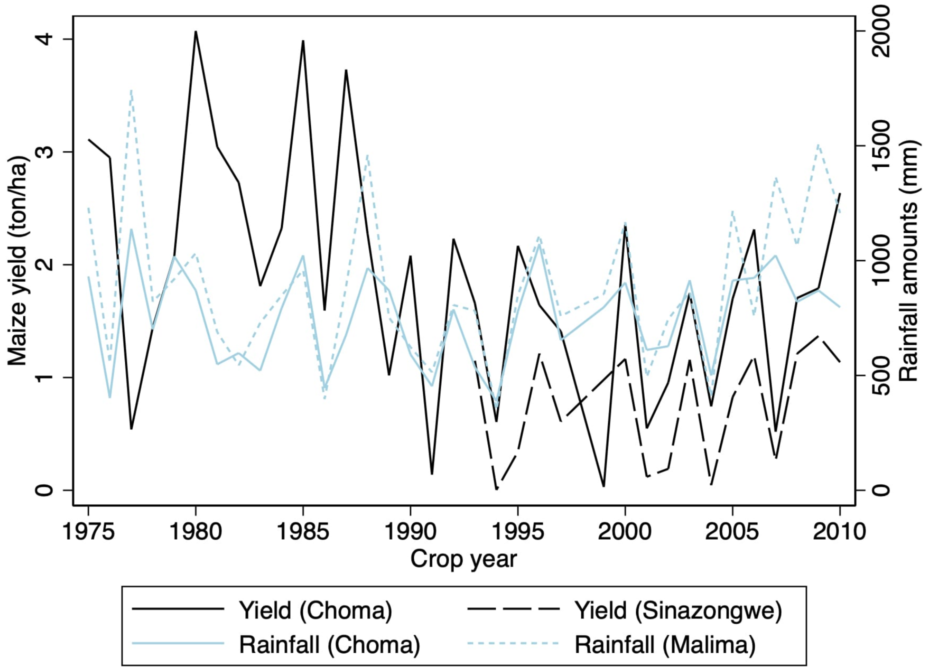


Figure 2: Rainfall and maize yield in Choma and Sinazongwe districts, 1975/76–2010/11.

Source: Crop forecast survey data from the Central Statistical Office; Rainfall data of Choma and Malima from the Choma Meteorological Station of Zambia Meteorological Department (Mochipapa) and the Malima irrigation site, respectively.

3 Data

3.1 Context

We use household survey data collected from maize farmers in Southern Province, Zambia. In the survey area, the main cropping season is the rainy season (November–April).[5] In this season, local farmers cultivate crops including maize, cotton, sweet potato, and groundnut under rainfed conditions. Given this nature, climate risk is the primary source to threaten their livelihoods. This is especially the case in Southern Province, known as the most drought-prone province in the country.

Figure 2 illustrates maize yields in Choma and Sinazongwe districts [6] and rainfall amounts recorded at the Choma meteorological station and Malima irrigation site (which is in Sinazongwe district) between 1975 and 2010. As evident in Figure 2, maize production co-moves with observed

[6]Our study sites are located in Sinazongwe district but one of the study site is close to Choma district. See Figure 3.

Table 3: Rainfall and maize yield in the study area, 1975/76–2010/11

	(1)	(2)	(3)	(4)	(5)
Rainfall					
Rainy season (100mm)	1.310**				
	(0.513)				
Rainy season, squared.	-0.080**				
	(0.035)				
Flowering season (100mm)		0.216**	0.747**		
		(0.092)	(0.346)		
Flowering season, squared.			-0.074*		
			(0.043)		
Planting season (100mm)		0.089	0.767		
		(0.119)	(0.501)		
Planting season, squared.			-0.106		
			(0.066)		
"Drought" in 11/12 contract				-0.573**	
				(0.279)	
"Flood" in 11/12 contract				-0.399	
				(0.425)	
"Drought" in 12/13 contract					-0.843***
					(0.218)
"Flood" in 12/13 contract					-0.973**
					(0.390)
Choma district	0.753***	0.807***	0.826***	0.811***	0.769***
	(0.207)	(0.216)	(0.192)	(0.228)	(0.204)
Linear time trend	-0.045***	-0.035**	-0.044**	-0.040***	-0.048***
	(0.014)	(0.016)	(0.018)	(0.015)	(0.012)
Constant	-3.115*	0.688	-0.926	1.977***	2.367***
	(1.738)	(0.506)	(0.877)	(0.450)	(0.395)
R-squared	0.470	0.412	0.477	0.387	0.546
Observations	51	51	51	51	51

Notes: The data sources were crop forecast survey data from the Central Statistical Office and rainfall data from the Choma Meteorological Station of Zambia Meteorological Department. The dependent variable is maize yield (mean = 1.52 tonnes/ha, Std.Dev. = 1.04). The sample covered the period between 1975/76 and 2010/11 for Choma district and between 1993/94 and 2010/11 for Sinazongwe district, with missing observations. "Drought" ("Flood") in the 2011/12 contract is a dummy variable that equals 1 if the total rainfall amount during the rainy season was below 600 mm (above 1000 mm), and 0 otherwise. "Drought" in the 2012/13 contract is a dummy variable that equals 1 if the total rainfall amount during January and February was below 280 mm, and 0 otherwise. "Flood" in the 2012/13 contract is a variable that equals 1 if the total rainfall amount in December was above 300 mm, and 0 otherwise. In pooling maize yield data from the Choma and Sinazongwe districts, OLS was used for the estimations. Heteroskedasticity-robust standard errors in parentheses. *$p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

precipitation patterns. To formally check their statistical relationship, Table 3 presents estimation results of regressing maize yields on rainfall and the other controls. As expected, total rainfall amounts during a rainy season have an explanatory power for maize harvests in the region: its impact has an inverted U-shape with a peak at about 820 mm (column 1).

Rather than the total rainfall amounts, the distribution of rainfall in the rainy season would matter more for maize cultivation. Local farmers specifically care about rainfall patterns in the two periods: the planting season (November–December) and flowering season (January–February). In this study, we focus more on flowering season rainfall as a determinant of farmers' maize yields for the following two reasons. First, historical data support the importance of flowering season rainfall relative to planting season rainfall. The linear specification in column 2 of Table 3 demonstrate that rainfall amounts during the flowering season have a significant relationship with maize yields unlike the planting season rainfall after controlling for flowering season rainfall. To account for the non-linear relationship, we add the squared terms of rainfall variables into the regression equation in column 3. We find that the estimated impacts of rainfall in the planting season are less precise than those for rainfall in the flowering season, although the coefficients are similar in magnitude. According to the regression results, yield predictions based on the flowering season rainfall would provide us with more reliable figures. Second, field observations also support the above view. Even if drought hits in the early stage of the rainy season, local farmers can re-plant early-maturing seed varieties to offset the loss. By contrast, insufficient rainfall during the flowering season limits crop growth, which directly leads to poor maize harvests.

3.2 Insurance sales and data collection

3.2.1 The sample

In this subsection, we describe our field surveys conducted in Southern Province. Our first survey dates back to the 2007/08 crop season. During the five years between 2007/08 and 2011/12, we have collected daily rainfall data at the plot-level and household information weekly from 48 households in five villages, in collaboration with the Zambia Agriculture Research Institute (ZARI). These five villages are in three different sites: sites A, B, and C (see Figure 3). Site A is a lower flat lakeside area with an elevation of 500–550 m. Site B is located in a middle escarpment area, and the elevation ranges between 750 and 1000 meters. Site C is in an upper terrace area on the Zambian plateau with an elevation of 1050–1100 m. Although these three sites are within a 15-km radius, their rainfall patterns are often distinct because of attitude-based differences. In each site, we randomly selected 16 households for the survey, providing a total sample of 48 households. We call this sample "old sample." After the end of the weekly household survey, we introduced rainfall index insurance to 100 households in November 2011 and 160 households in 2012/13 and 2013/14 crop years. These target households include the old sample.

We mainly focus on 48 households in the old sample for two reasons. First, we collected plot level rainfall data only from these farm households. The plot-level rainfall dataset is crucial for this study to estimate household-specific basis risk in our empirical analysis. Second, the long-term relationships with the research team and ZARI might alleviate problems associated with a lack of trust in insurance sales, which is one of the potential barriers to insurance demand (Giné and Yang, 2009). [7]

3.2.2 Insurance products

Under the collaboration with the ZARI and the Zambia Meteorological Department, we introduced a new weather index insurance product in November 2011. The insurance payout is conditional only on an objective indicator (i.e., rainfall at a pre-specified weather observation point) and designed to cover drought and flood events. The premium was set at ZMW 5 (about US$ 1.1) per unit.[8] As the average wage for agricultural casual labor was approximately ZMW 10 per day at the survey time, the premium was set and fixed at a level that was affordable for local farmers.

To set the actuarially fair premium rate, we used the estimation results based on historical data in Table 3 and designed weather index insurance contracts for each year as follows. In the 2011/12 crop year, the defined condition under which a farmer would receive an insurance payout was if the total rainfall recorded at the Choma weather station is either less than 600 mm or more than 1000

[7] As illustrated in Appendix Table A3, we observe that the average insurance uptake units were higher for the households in the old sample compared to the counterparts in the new one.

[8] ZMW = Zambian Kwacha. In November 2011, the exchange rate was 5.2 ZMW/US$.

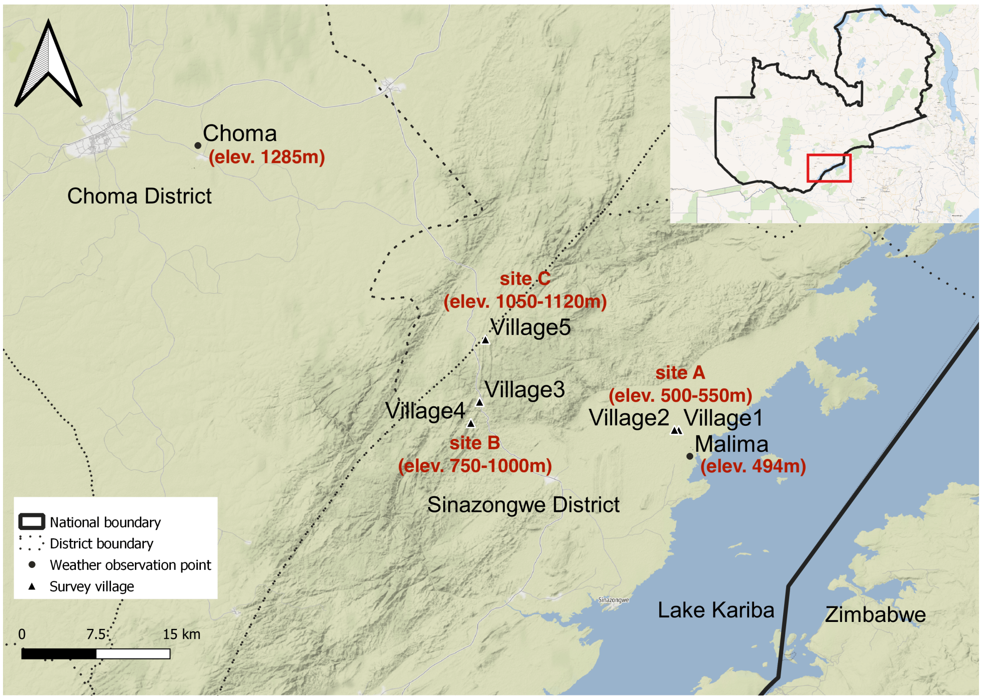


Figure 3: Map of study site and weather observation points

Notes: For Choma, the weather observation point is the Choma Meteorological Station of the Zambia Meteorological Department. For Malima, the weather observation point is the Malima irrigation site.

mm during the rainy season. Column 4 of Table 3 confirms that the "drought" defined in the first year contract decreased maize yield in the corresponding crop year by 0.57 tonnes/ha.[9] Referring to the historical rainfall data, we set 20% ($q = 1/5$) as the premium rate, and thus, the insurance payout per unit was ZMW 25 ($= 5/q$). For the insurance sales in the 2012/13 crop year, we used a more precise index: farmers would receive the insurance payout if rainfall in the flowering season (January and February) at the Choma weather station is less than 280 mm or rainfall in December is more than 300 mm. Both specified events are found to reduce maize yield by 0.84 tonnes/ha and by 0.97 tonnes/ha, respectively (column 5 of Table 3). In the last year, we referred to rainfall recorded at the Malima irrigation site.[10] This new observation site is much closer to the study villages (Figure 3). Reflecting local farmers' repeated requests for the cover of rainfall shocks in the planting season, we used rainfall amounts in the planting season (November and December) as the

[9] By contrast, the defined "flood" does not have a significant effect. Nevertheless, we covered flood events in the insurance contract to reflect the demand among local farmers. We further discuss this point in Section 3.2.5.

[10] In the first two years, we did not realize the availability of rainfall records collected by the local development project at the Malima irrigation site.

index in the insurance policy: farmers would be compensated when total rainfall recorded at the Malima irrigation site is less than 214 mm or exceeds 800 mm in November and December.[11] The historical rainfall data suggest $q = 1/3$ for payout conditions in the last two years, and thus, the insurance payout per unit was ZMW 15.

Figure 3 depicts the locations of the survey sites, the Choma weather station, and the Malima irrigation site. All study sites were located far from Choma while sites B and C were relatively closer. In particular, those two sites were about 35–40 km away from the station, whereas site A was as far as 55 km. Therefore, we expect spatial basis risk to be higher in site A under the insurance contract in the first two years. By contrast, the Malima irrigation site was close to all survey villages, especially site A.

3.2.3 Door-to-door insurance marketing[12]

We sold the insurance product every early November (i.e., before the start of the rainy season) during the three crop years, 2011/12–2013/14. In each year, our trained enumerators visited the survey households 10 days before the insurance sale. First, they explained the insurance contract to the household head, left a copy of a leaflet visually explaining it, and informed the head that insurance sales would occur at a designated place–usually the village head's place–approximately 10 days later. In addition, the enumerators told the survey participants that if they were interested in the insurance, they should bring enough money for the purchase.

In the door-to-door visits, the enumerators also conducted interviews with the household heads to collect information relevant to insurance demand (e.g., subjective perceptions of weather risk, arithmetic skills, and financial literacy) as well as basic demographic information. The interview included the Binswanger (1980) style lottery game to elicit attitudes towards risk among the household heads. The game was incentivized with actual winnings. In the game, the enumerator showed a household head six alternatives with different rewards, as listed in Table 4, and explained that winnings were dependent on the result of a coin tossed by the enumerator and would be paid on the

[11]The change in the reference index was for this practical reason. Using historical rainfall data from the Malima irrigation site, Appendix Table A1 replicates Table 3. Estimation results confirm the importance of rainfall for both seasons. The design of the last year's contract prioritized perceptions among farmers over statistical evidence.

[12]This and the next subsections draw heavily on Miura and Sakurai (2015).

Table 4: Design of Binswanger-style lottery

Option	Heads Low Pay-off	Tails High Pay-off	Risk-Aversion Class	Coefficient of Partial Risk-aversion
a	ZMW 5	ZMW 5	Extreme	$3.93 \leq \sigma \leq \infty$
b	ZMW 4	ZMW 12	Severe	$1.00 \leq \sigma \leq 3.93$
c	ZMW 3	ZMW 16	Intermediate	$0.56 \leq \sigma \leq 1.00$
d	ZMW 2	ZMW 19	Moderate	$0.27 \leq \sigma \leq 0.56$
e	ZMW 1	ZMW 21	Slight-to-neutral	$0.00 \leq \sigma \leq 0.27$
f	ZMW 0	ZMW 22	Neutral-to-negative	$-\infty \leq \sigma \leq 0.00$

Notes: ZMW 1 = about US$ 0.22

insurance sales day. Based on the choice, we categorized the respondents into risk-aversion classes (Table 4).

3.2.4 Insurance sales and result reports

After about 10 days of door-to-door visits, we held a sales day for each village. First, we asked each household head how many insurance units they wanted to purchase. Then, each participant paid money per their demand. At that time, they could use the rewards of the Binswanger game for the payment, if he or she had won. This allows us to test the role of liquidity constraints on insurance demand, as proposed in the previous literature (e.g., Cole et al., 2013; Giné et al., 2008).

After the end of the contract period every year, the research team and ZARI personnel visited the study villages to inform the survey participants of whether or not an insurance payout would happen. The rainfall index did not satisfy the payout conditions in all the years.[13] Thus, no payouts were made throughout the research period.[14]

[13] In the first year, the total amount of rainfall during the 2011/12 rainy season was 711.7 mm at the Choma weather station. In the second year, the total rainfall amount at the Choma weather station during December 2012 was 200 mm, and that during the flowering season was 510.7 mm. In the last year, the total rainfall amount in November and December 2013 was 379.5 mm at the Malima irrigation site.

[14] At the end of the multi-year research project (May 2014), we returned to the surveyed farmers their premiums, based on our past sales records. During the intervention, this was kept a secret from the respondents.

Table 5: Rainfall at plots and weather reference points during the flowering season

	2007/08	2008/09	2009/10	2010/11	2011/12
Site A (16 households)	521.8	458.4	912.0	424.5	322.8
	(13.9)	(69.5)	(61.0)	(24.8)	(24.9)
Site B (16 households)	525.7	559.9	843.3	625.2	375.1
	(19.6)	(25.3)	(123.3)	(63.5)	(30.7)
Site C (16 households)	494.4	603.2	721.2	442.8	333.0
	(32.0)	(28.2)	(94.8)	(31.0)	(28.2)
Total (48 households)	514.0	540.5	825.5	497.5	343.6
	(26.6)	(75.9)	(123.5)	(100.9)	(35.7)
Choma meteorological station	394.4	313.3	369.1	334.5	252.4
Malima irrigation site	366.3	459.9	959.1	670.0	286.1

Notes: The numbers represent the average total rainfall in January and February in millimeters. Standard deviations in parentheses.

3.2.5 Plot-level rainfall data

The final data used for our empirical analysis are plot-level rainfall records. In November 2007, we installed automatic rainfall gauges and loggers at the representative plot for the 48 households in the old sample. Since then, we recorded daily rainfall amounts at the plot-level for each surveyed household until April 2012. We used the collected field-level rainfall data to compare with the rainfall index observed at the weather reference points when estimating the probabilities of both drought and false–negative for each household.

Table 5 summarizes rainfall amounts during the flowering season at each site and the two weather observation points. Rainfall data indicates non-negligible differences in precipitation among the locations within a small area, suggesting the salience of spatial basis risk. Although heavy rainfall and floods are rare in the local area, farmers experienced them in the 2009/10 crop year. This is the main reason we covered floods as well as droughts in our insurance contracts. However, our empirical analysis only considers drought risks as flood risks are relatively insignificant in the local context and the incorporation complicates the analysis without generating additional implications.

Table 6: Descriptive statistics

	2011/12		2012/13		2013/14	
	Mean	Std.Dev.	Mean	Std.Dev.	Mean	Std.Dev.
Units of insurance takeup (#)	2.929	(2.224)	2.614	(1.956)	2.405	(1.231)
Expected output in normal year (tonne)	.	(.)	1.640	(1.025)	.	(.)
Expected output in drought year (tonne)	.	(.)	0.527	(0.351)	.	(.)
Expected loss (tonne)	.	(.)	1.113	(0.744)	.	(.)
RA – Extreme (dummy)	0.381	(0.492)	0.068	(0.255)	0.095	(0.297)
RA – Severe (dummy)	0.119	(0.328)	0.250	(0.438)	0.238	(0.431)
RA – Intermediate (dummy)	0.095	(0.297)	0.205	(0.408)	0.143	(0.354)
RA – Moderate (dummy)	0.071	(0.261)	0.227	(0.424)	0.310	(0.468)
RA – Slight (dummy)	0.167	(0.377)	0.114	(0.321)	0.119	(0.328)
RA – Neutral (dummy)	0.167	(0.377)	0.136	(0.347)	0.095	(0.297)
Cash reward (ZMW)	7.929	(6.475)	11.682	(7.706)	12.857	(7.166)
Observations	42		44		42	

Notes: ZMW 1 = about US$ 0.22. RA denotes risk-aversion class.

3.3 Descriptive statistics

Table 6 presents descriptive statistics of insurance sales results and key variables from the household interviews by survey year. Appendix Table A4 reports the summary statistics of purchased units by site as well as that of the old sample. In each year, most farmers purchased more than two units of the insurance contract. The average insurance uptake was the highest in the first year.

The expected maize outputs in normal and drought years were asked only during the household interview in November 2012. On average, farmers expect to lose 1.113 tons (67.9% of the normal harvest) of maize harvests in a drought year. We regard this expected loss as time-invariant and use this value for our estimations. In economic terms, the expected loss of 1.113 tons of maize is equivalent to approximately ZMW 835.[15] This implies that one unit of the insurance product covered s 2.99% (25/835) in 2011/12 and 1.80% (15/835) in 2012/13–2013/14 of the expected loss (cover rate per unit). Thus, average farmers covered only 8.77%, 4.70%, and 4.32% of their expected loss with the purchased rainfall index insurance products. These low cover rates are consistent with that reported in the literature (e.g., Cai and Song, 2017; Dercon et al., 2014; Jensen et al., 2018).

[15] According to the 2012/13 household survey, mode and median maize prices for household consumption were both ZMW 0.75 per kg. While only a few farmers sold maize in that year, most reported the gateway price as a range of

Table 6 also presents the results of the risk game. We categorize the household heads into one of the risk attitude groups based on the choice in the Binswanger-style lottery. Although risk preferences should be constant, we find that the shares of each risk-aversion class have changed across the survey years. Such changes can have two possible reasons. The first is learning. The majority chose the safest option (i.e., always receive ZMW 5) only in the first year. Farmers could have learned how the lottery game works and took some risks in the last two years. Second, farmers may have changed their risk attitudes over time according to their experiences (cf, Cai and Song, 2017). As more respondents chose risky options, the average winnings amounts were higher in the last two years than to the first year.

4 Empirical framework

4.1 Econometric specification

During the three survey years, we sold weather index insurance in different designs (i.e., rainfall index, threshold, and reference location). This implies that a household faced different basis risks across time. This within-household variation allows us to employ the household fixed effect model to test whether a farmer responds to an increase (or a decrease) in the basis risk. Following the theoretical model, our main specification is:

$$\alpha_{it} = \beta_0 + \beta_1(r_{it} \times \mathrm{loss}_{it}) + \mu_i + \mu_t + \varepsilon_{it}, \tag{3}$$

where ε_{it} is a mean zero error term that may be correlated within the household i; α_{it} is the number of insurance contracts that household i purchased in year t; loss_{it} is a difference in expected outputs between normal and drought years, an empirical counterpart to $y_i - \underline{y}_i$ in our theoretical model; r_{it} represents basis risk or probability of false–negative estimated in the next subsection using rainfall data at plot-level and weather observation locations; μ_i and μ_t are fixed effects for household i and year t, respectively; and β_1 is interpreted as a change in the purchased unit of contracts in

ZMW 0.6–0.8 per kg. We presume that loss of 1000 kg (1 ton) is equivalent to ZMW 750 in economic terms and estimate the average economic loss in drought year as ZMW 835($\approx$750$\times$1.113).

response to one percentage point change in the basis risk. When $\beta_1 < 0$ is confirmed, the basis risk is found to decline insurance demand as predicted by our theoretical model. This fixed effect estimate exploiting within-household variations of the basis risk enables us to estimate a precise and unbiased parameter that represents the impact of the basis risk on farmers' behavior of actual insurance purchase.

In our setting, the probability of plot-level drought p is constant across years for household i. In addition, the probability of insurance payment q is constant across households in year t. However, the model predicts that these probability components are also determinants of the insurance demand (see equation (2)). To investigate the relationship between a set of two probabilities (p, r) and the insurance uptake α, we run the following regression with the pooled data combining the three years. Building on the optimal uptake unit in equation (2), we specify the regression model as:

$$\alpha_{it} = \gamma_0 + \gamma_1(r_{it} \times \text{loss}_{it}) + \gamma_2(p_i \times \text{loss}_{it}) + \mu_t + \varepsilon_{it}, \tag{4}$$

where ε_{it} is an error term. The empirical evidence is consistent with model predictions when $\gamma_1 < 0$ and $\gamma_2 > 0$ are confirmed. Subsection 4.2 describes how we estimate the key variables in our econometric specifications (3) and (4): p_i and r_i.

4.2 Estimating probabilities

To estimate the regression equations (3) and (4), probabilities of drought and false–negative (p_i, r_i) should be quantified. The estimation of p_i, r_i builds upon key assumptions on maize production and probability distributions of rainfall. Let R_i, R_c, R_m be the observed rainfall at household i's plot, the Choma weather station, and the Malima irrigation site, respectively. We introduce the following assumptions:

- Assumption 1: Farmers incur crop loss because of "drought" when R_i during the flowering season (January–February) is 280 mm or less.
- Assumption 2: R_i follows a truncated normal distribution.
- Assumption 3: (R_i, R_c) and (R_i, R_m) follow a truncated multivariate normal distribution.

Assumption 1 specifies flowering season rainfall as a key determinant of maize harvests. The statistically significant results in Table 3 justify Assumption 1 in the research context. Assumption 1 implies no product basis risk (or at least constant among farmers) when the index refers to rainfall during the flowering season with 280 mm as the drought threshold. Under assumption 1, a source of basis risk was spatial basis risk alone in the 2012/13 insurance contract while farmers face both spatial and product basis risks in the 2011/12 and 2013/14 insurance contracts. Farmers faced a similar spatial basis risk under the 2011/12 and 2012/13 contracts because Choma was the reference weather station in both years. In the 2013/14 contract, farmers were exposed to distinct spatial basis risk because of a shift of the weather reference location to the Malima irrigation site. They may also have different product basis risks in the last year because the index was changed to rainfall during the planting season.

Assumption 2 about the probability distribution of rainfall emerges from historical rainfall data at the two weather reference points. Appendix Figure A2 presents the historical distribution of the flowering season rainfall at the Choma weather station during 1949/50–2012/13, which seems to support Assumption 2. Various statistical tests also validate the normality assumption (see Appendix Table A2 for test results). We also confirm a similar rainfall pattern in rainfall data from the Malima irrigation site. As precipitation is non-negative, we impose an additional assumption that the distribution is truncated at zero. Given these, we extend the assumption of truncated normality to the joint distribution of field-level rainfall and indexed rainfall (Assumption 3).

Under assumptions 1 and 2, we can derive a probability of drought at the plot:

$$p_i = \int_0^{280} f_{R_i^{1-2}}(s)ds,$$

where $f_{R_i^{1-2}}$ is a probability density function for the truncated normal distribution of the plot-level rainfall in the flowering season (January–February). The first and second moments of the distribution are estimated by adjusting the five-year plot-level rainfall data based on historical rainfall data from the Choma meteorological station, 1949/50–2011/12. The estimated p_i is constant across the survey years, 2011/12–2013/14. By contrast, the probability of false–negative because of the spatial and product basis risks differs by survey year because of changes in the insurance payout

Table 7: Descriptive statistics of estimated probabilities

	2011/12		2012/13		2013/14	
	Mean	Std.Dev.	Mean	Std.Dev.	Mean	Std.Dev.
p, probability of drought	0.130	(0.043)	0.130	(0.043)	0.130	(0.043)
Site A (16 households)	0.173	(0.016)	0.173	(0.016)	0.173	(0.016)
Site B (16 households)	0.095	(0.042)	0.095	(0.042)	0.095	(0.042)
Site C (16 households)	0.123	(0.022)	0.123	(0.022)	0.123	(0.022)
r, probability of false-negative	0.078	(0.040)	0.060	(0.036)	0.094	(0.029)
Site A (16 households)	0.124	(0.014)	0.103	(0.013)	0.121	(0.010)
Site B (16 households)	0.050	(0.030)	0.036	(0.024)	0.072	(0.031)
Site C (16 households)	0.060	(0.019)	0.041	(0.016)	0.091	(0.018)
Observations	48		48		48	

conditions:

$$2011 : r_i = \int_0^{280} \int_{600}^{\infty} \zeta_{R_i^{1-2}, R_c^{11-4}}(s,t)dsdt,$$
$$2012 : r_i = \int_0^{280} \int_{280}^{\infty} \zeta_{R_i^{1-2}, R_c^{1-2}}(s,t)dsdt,$$
$$2013 : r_i = \int_0^{280} \int_{214}^{\infty} \zeta_{R_i^{1-2}, R_m^{11-12}}(s,t)dsdt.$$

As for the 2011/12 and 2012/13 years, $\zeta_{R_i^{1-2}, R_c^{11-4}}$ and $\zeta_{R_i^{1-2}, R_c^{1-2}}$ are probability density functions of a truncated binormal distribution of plot-level rainfall in the flowering season and observed total rainfall during the rainy season (November–April) and flowering season rainfall, respectively, at the Choma meteorological station. For the 2013/14 crop year, $\zeta_{R_i^{1-2}, R_m^{11-12}}$ is the probability density function that jointly draws flowering season rainfall at the plot-level and planting season (November–December) rainfall at the Malima irrigation site. Moments and correlations for the joint probability density function are estimated base on rainfall data from household plots and reference sites.[16]

Table 7 presents descriptive statistics of a set of estimated probabilities and Figure 4 depicts

[16]For example, mean, standard deviation (Std.Dev.), and correlation to derive a probability of false–negative for the 2011/12 crop year are estimated as follows: (i) Mean and Std. Dev. of Choma rainfall originates from the sample average and Std.Dev. of the total rainfall during the rainy season (November–April) between 1949/50 and 2011/12; (ii) Mean and Std.Dev. of plot i's rainfall is based on the sample average and Std.Dev. of the flowering season rainfall during the sample period adjusted by historical data on the flowering season rainfall at the Choma weather station (e.g., $\hat{E}(R_i) = \left(\sum_{t=2007}^{2011} R_{it}^{1-2} / \sum_{t=2007}^{2011} R_{ct}^{1-2}\right) \times \frac{1}{63} \sum_{t=1949}^{2011} R_{ct}^{1-2}$); (iii) Correlation is computed with total rainfall during the rainy season at the Choma weather station and flowering season rainfall at plot i.

their distributions. Farmers in site A faced higher risk of drought and basis risks in all the years than those in other sites. The probability of false–negative is higher in 2011/12 compared to that in 2012/13. This difference would reflect the product basis risk because of the change in insurance contract designs. Although we expected the basis risk to be smaller because of the closer reference location in the last year's contract, the estimated basis risk turns out to be relatively high in the year. This unexpected result may imply a higher product basis risk in the last year. Although Site A is farthest from Choma and closest to Malima, the farmers faced higher basis risk in all years. This counterintuitive result may reflect higher exposure of farmers to drought risks in Site A. Overall, the estimated basis risk changed across the years for the survey households in all sites. This within-household variation encourages our fixed effect specification in equation (3). Finally, we confirm that for most farmers, the estimated probabilities (p_i, r_i) and a probability of payout ($q = 1/5$ in 2011/12 and $q = 1/3$ in 2012/13–2013/14) satisfy the theoretical assumption imposed on the probability structure (i.e., $p - q < r < p(1 - q)$).

5 Results

In this section, we first present the estimation results of our baseline fixed-effect model, followed by the test of model predictions. Based on the empirical results, we discuss the economic impact of the basis risk. Finally, we discuss other possible constraints of insurance demand and their interplay with the basis risk.

5.1 Baseline results

Table 8 presents our baseline regression results. The dependent variable is the purchased unit of insurance contracts. Our estimation results confirm that farmers adjust their purchased units in response to a change in basis risk. As the model predicts, we detect the significant effect of within-household variation in basis risk only when interacting with expected loss (Columns (1) vs. (2)). Columns (3) and (4) confirm the robustness of this empirical pattern to the inclusion of dummies

[17] In Appendix Table A7, we also check its robustness to bootstrap standard errors.

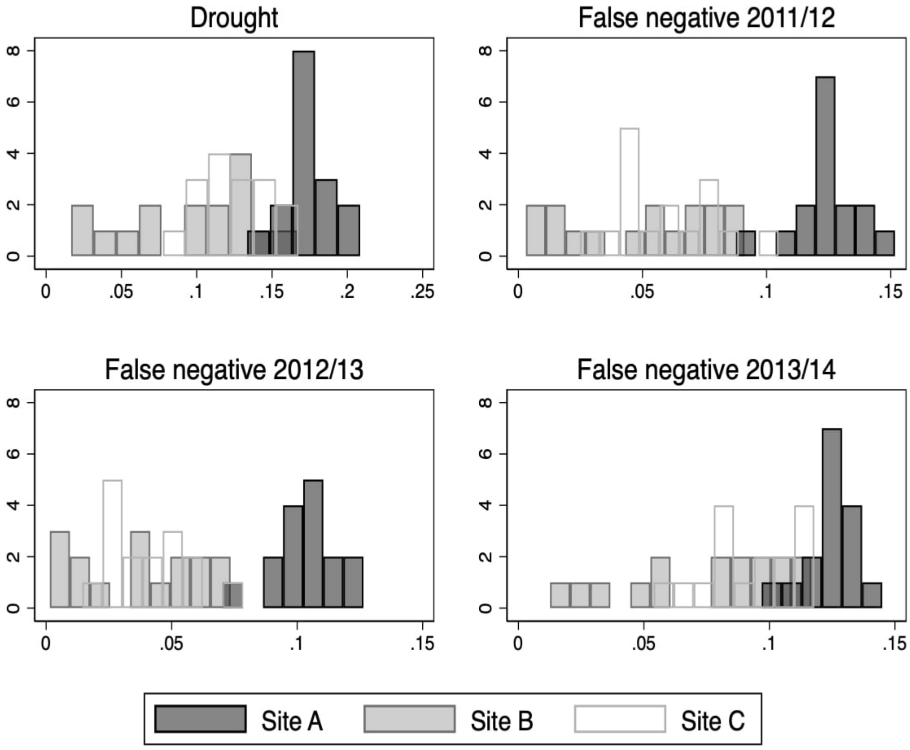


Figure 4: Estimated probabilities of drought and false–negative

for risk-aversion classes.[17] In addition, we do not find that the cash reward–the byproduct of the Binswanger-style lottery–is a significant determinant of the purchase. The null effect of immediate cash would be natural given that we set the premium to ZMW 5 (approximately US$1) that was affordable even for worse-off farmers.

According to the result in column (3), the average household with 1.11 tonnes of maize loss due to drought increases insurance uptake by 0.925 units in response to a 5 p.p. decrease in the probability of false–negative, p. Given that the average of estimated p is 13.0%, this impact of the basis risk on insurance demand is limited. When this impact is evaluated in terms of the cover rate, we also find the modest welfare gain of reducing the basis risk. In each year, on average, farmers covered 8.77% (2011/12), 4.70% (2012/13), and 4.32% (2013/14) of their expected loss with our index insurance product. A 5 p.p. decrease in basis risk, for example, could have increased their cover rate by 2.77 p.p. or 31.58% (2011/12), 1.66 p.p. or 35.39% (2012/13), and 1.66 or 38.46%

Table 8: Effect of the basis risk on insurance demand: Fixed effect estimates

	(1)	(2)	(3)	(4)	(5)
r × Loss	-14.354**	-14.054*	-16.217**	-16.107**	
	(6.913)	(7.273)	(7.751)	(8.005)	
r		-4.801		-1.844	-19.765
		(18.086)		(15.047)	(16.761)
RA–Extreme/Severe			-0.582	-0.592	-0.518
			(0.403)	(0.411)	(0.387)
RA–Intermediate			0.519	0.511	0.349
			(0.531)	(0.537)	(0.411)
RA–Slight			0.900*	0.891*	0.871*
			(0.491)	(0.492)	(0.451)
RA–Neutral			0.288	0.280	0.151
			(0.483)	(0.505)	(0.464)
Cash reward			0.026	0.026	0.023
			(0.029)	(0.029)	(0.035)
Year = 2012	-1.188**	-1.257*	-1.550**	-1.577**	-0.894
	(0.589)	(0.667)	(0.689)	(0.729)	(0.579)
Year = 2013	-0.231	-0.148	-0.401	-0.369	-0.422
	(0.382)	(0.421)	(0.431)	(0.416)	(0.443)
Constant	4.132***	4.493***	4.149***	4.299***	4.442**
	(0.658)	(1.517)	(0.837)	(1.436)	(1.670)
R-squared	0.678	0.678	0.739	0.739	0.680
Observations	126	126	126	126	126

Notes: The dependent variable is the number of insurance contracts that a household purchased. RA denotes risk-aversion class and RA–Moderate is used as a reference. Robust standard errors clustered by household in parentheses. *$p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

(2013/14). Even after this increase, the expected cover rate remains low: 11.54% (2011/12), 6.36% (2012/13), and 5.98% (2013/14). As a policy intervention to decrease the basis risk by 5 p.p. would be very costly in practice, our results imply that minimizing the basis risk would not yield enough economic benefits to offset the associated costs.

5.2 Testing the model predictions

To test the theoretical predictions in equation (2), we also run the OLS regression with the pooled data. Table 9 presents the estimation results. Across all the specifications, the coefficients on

Table 9: Test of the model predictions: Pooled OLS estimates

	(1)	(2)	(3)	(4)	(5)	(6)
r × Loss	-33.962***	-37.150***	-34.268***	-38.339**	-34.052***	-38.318**
	(11.639)	(12.685)	(10.757)	(14.405)	(11.375)	(14.416)
p × Loss	26.564***	38.708***	26.689***	36.064**	26.315***	35.761**
	(8.323)	(13.429)	(7.151)	(15.132)	(7.497)	(14.521)
r		20.311		19.060		20.243
		(22.019)		(22.054)		(21.264)
p		-32.940		-26.214		-28.074
		(23.533)		(24.110)		(22.491)
Loss		-1.502		-1.010		-0.971
		(1.698)		(1.613)		(1.543)
RA–Extreme/Severe			-1.351***	-1.314**	-1.355***	-1.290***
			(0.499)	(0.498)	(0.485)	(0.472)
RA–Intermediate			0.175	0.190	0.143	0.223
			(0.474)	(0.452)	(0.421)	(0.393)
RA–Slight			-0.073	-0.092	-0.093	-0.093
			(0.648)	(0.651)	(0.671)	(0.677)
RA–Neutral			-0.176	-0.150	-0.150	-0.081
			(0.477)	(0.475)	(0.450)	(0.435)
Cash reward			0.029	0.026	0.028	0.022
			(0.027)	(0.026)	(0.028)	(0.029)
Year = 2012	-1.015**	-0.718	-1.430***	-1.154**	-1.436***	-1.122**
	(0.425)	(0.525)	(0.481)	(0.557)	(0.480)	(0.535)
Year = 2013	0.156	-0.116	-0.279	-0.489	-0.289	-0.484
	(0.309)	(0.368)	(0.377)	(0.408)	(0.386)	(0.415)
Constant	1.975***	4.946**	2.475***	4.575**	2.830***	4.856**
	(0.368)	(2.068)	(0.538)	(1.951)	(0.609)	(2.031)
Household characteristics	No	No	No	No	Yes	Yes
R-squared	0.219	0.246	0.374	0.388	0.380	0.396
Observations	126	126	126	126	126	126

Notes: The dependent variable is the number of insurance contracts that a household purchased. RA denotes risk-aversion class and RA–Moderate is used as a reference. Heteroskedasticity-robust standard errors clustered by household in parentheses. Household characteristics include the household head's sex, age, and years of schooling. *$p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

the interaction terms between a set of probabilities (p, r) and the expected loss are statistically significant in the expected direction. These significant results remain even after controlling for risk preferences and basic household characteristics. One may concern the assumption of drought in the flowering season as a determinant of poor harvest (Assumption 1 in Section 4). As the relationship of maize yield with rainfall at the Choma weather station provides empirical support (Table 3), this assumption may not hold for agricultural production in places distant from Choma (e.g., villages in Site A). To address this concern, Appendix Table A5 presents the same result except for using the sample restricted to Sites B and C. This additional regression produces a similar empirical pattern to that in Table 9, although the result is less precise in some specifications because of a small sample size. Overall, the results in Tables 9 and A5 validate our proposed theoretical model and further justify the baseline specification.

Before proceeding, it would be worthwhile to discuss other empirical variables. As predicted in the approximated solution of insurance demand, risk preferences do not appropriately predict the purchased units of the contract, except for the Extreme/Sever risk-aversion class. This result is in line with the predictions from our numerical exercise that demonstrates a drop in insurance demand specifically for the most risk-averse farmers because of the basis risk (Panels 1a and 1b). Finally, Tables 8 and 9 indicate that insurance uptake is significantly lower in the 2012/13 crop year than in the 2011/12 year. The first year had no insurance payout even though some farmers complained that they had experienced drought. Their expected benefits of the insurance product might have been updated in a negative direction (Jensen et al., 2018). Such behavioral factors may lead to the decrease in insurance demand in the second year.

5.3 Discussion

We found a statistically significant association between the basis risk and farmers' insurance purchase behavior. The results also suggest that the economic significance of the basis risk is modest given the high potential costs associated with efforts to reduce the basis risk (e.g., placing more weather stations). The evidence suggests that potential demands for index insurance products with the minimum the basis risk would be limited under the current economic conditions. This takeaway from our study is in line with the recent experience of IBLI sales in northern Kenya. To minimize the spatial basis risk, the IBLI uses the Normalized Difference Vegetation Index observed by satellite platforms as the index (cf, Jensen et al., 2016; Takahashi et al., 2020). Despite this, Jensen et al.

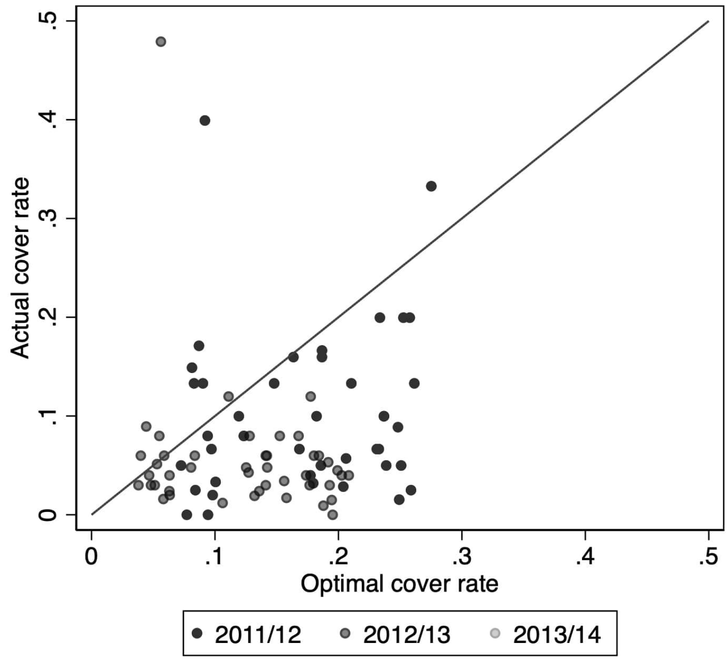


Figure 5: Actual vs. optimal cover rates

Notes: An optimal cover rate is computed by substituting a set of estimated probabilities (p, r) and fixed q into equation (2). Following the assumption, samples with the probabilities that do not satisfy $p - r < r < p(1 - q)$ are excluded from the figure.

(2018) report that the overall uptake rate was 16% and the average cover rate for those who joined the scheme was 24%.[18]

In our insurance sales in Zambia, we observed low insurance demand over the years. The actual cover rates were much lower than the optimal cover rate predicted in our theoretical model (equation 2). Figure 5 depicts the optimal and actual cover rates among the surveyed households. The figure shows that the majority were located below the 45-degree line: they purchased less than the optimal unit of the insurance contracts predicted by the theoretical model. This result suggests that constraints other than the basis risk hinder insurance demand among maize farmers in our study site.

To investigate the potential barriers to spurring demand for index insurance, we discuss the

[18] An ongoing study by Afshar et al. (2021) is another interesting case in this regard. Their study in India aims to minimize the basis risk by monitoring crop phenology with satellite remote sensing and applying crop simulation models. Their results provide additional evidence in this context.

importance of the basis risk relative to other factors pointed out in the literature. In particular, we discuss the role of financial literacy, informal risk-sharing, and experience of drought. To facilitate the discussion, we run the regression equation (4) again with additional empirical variables. Given the data limitations, we only use data from the first two years for this exercise. Appendix Table A6 reports estimation results.

First, the literature acknowledges the important role of financial literacy in insurance purchases. For example, Gaurav et al. (2011) report that the provision of financial literacy and insurance education modules increased the uptake rate of rainfall insurance by 8% to 16% in India. Some studies also discuss the effectiveness of experimental games to help individuals with limited formal education understand the general concept of insurance (e.g., Cai and Song, 2017; Jensen et al., 2018; Janzen et al., 2020). In the RCT-based study in Kenya, Janzen et al. (2020) randomly provided an opportunity to play experimental games and asked farmers' willingness to pay for an index insurance product with the differentiated basis risk. They found farmers are sensitive to basis risk but their sensitivity was not affected by the intervention of the experimental game. In our household survey, we asked several questions to measure farmers' comprehension of our insurance contract and general arithmetic skills. Appendix Table A6 illustrates the evidence of their effects on the insurance demand. However, we find no evidence that knowledgeable farmers are more responsive to the basis risk, which is consistent with Janzen et al. (2020).

Second, we consider the role of informal risk-sharing. Farm households in developing countries share idiosyncratic risks with their relatives, friends, and neighbors. On the one hand, this pre-existing informal risk-sharing scheme may reduce the incentive to join formal insurance schemes (Arnott and Stiglitz, 1991). On the other hand, Mobarak and Rosenzweig (2012) argues that when idiosyncratic risks are informally shared across villages, farmers may mitigate basis risk, which differs spatially, through the informal risk-sharing arrangements. In line with this, Dercon et al. (2014) theoretically demonstrate that informal risk sharing and index insurance are complements, whereas indemnity-based insurance is substitutable. However, this argument may not be valid when households within the informal risk-sharing group face similar basis risks. To observe the relationship, we add the number of people whom the household could call upon in times of need as an empirical proxy for access to informal risk-sharing to equation (4). Our estimation results indicate that this variable does not have the explanatory power for insurance demand (Appendix Table A6). As this proxy may not suitably capture the strength of informal risk coping, this point needs to be further investigated.

Lastly, we discuss the role of recent experiences of weather shocks in insurance demand. We exploited plot-level rainfall data to estimate probabilities of drought and false–negative for each household to test theoretical predictions. As such, our empirical analysis and its results build upon the validity of objective probabilities. However, local farmers may formulate their decisions based on their subjective expectations about the likelihood of weather events, and thus, insurance payout.[19] In this context, the previous literature examined the impact of the recent crop loss and/or insurance payout. For instance, Cai and Song (2017) demonstrate that the experience of crop loss in the previous year adversely affected insurance demand, albeit fostering insurance literacy through playing the hypothetical game had substantially higher impacts. We also test this by adding the drought experienced in the last year into the right-hand side of equation (4). In contrast to what the previous studies reported, our estimation results in Appendix Table A6 do not reflect any significant effect of drought experiences on insurance demand. However, it is early to undervalue farmers' perceptions in their decision-making. For instance, few studies investigate the role of perceptions about both extreme weather events and basis risk in decision-making. This intriguing point is another gap to fill in future work.

6 Conclusion

In this study, we tested whether the basis risk is a barrier to demand for index insurance in rural Zambia. Using unique rainfall data collected at the households' plots for five years, we estimated the theoretical probability of no insurance payout when crop loss happens for each household. We then related this direct measure for the basis risk to actual demand for the rainfall index insurance contract introduced in the survey area. Our theoretical model on the farmers' optimal behavior of their insurance purchase motivated the econometric specifications. Exploiting contract design

[19] As discussed, Jensen et al. (2016, 2018) report a negative impact of experiencing "basis error", which is estimated as the household-level basis risk, on the IBLI demand.

changes during the survey period and within-household variations in the basis risk, we provided empirical evidence for the negative impact of the basis risk on household demand for rainfall index insurance contracts. Moreover, the cross-sectional estimation results validated our theoretical model. The provision of such direct evidence on the relationship of basis risk with insurance demand is our first contribution to the literature.

Another important lesson from this study is that, despite its statistical significance, the economic impact of the basis risk on insurance demand is modest. For example, our estimation results revealed that a 5 p.p. reduction in the probability of false–negative (more than a half of the average) would increase the cover rate of expected crop loss through insurance by 1.66–2.77 p.p. (32–38%) at most. This result implies that even if we introduce an insurance product designed to minimize the basis risk, farm households will cover a small fraction of their potential loss due to weather shocks. Therefore, the policy design meant to minimize the basis risk would never be cost-effective because it would not yield enough welfare impacts to offset the high associated costs.

Our results build on a small sample from a particular area in Africa. To check the external validity of our findings, further data collection that provides us with an opportunity to quantify household-specific basis risk is required. In addition, the interaction of basis risk with other potential impediments (e.g., informal risk-sharing arrangements) remains unsolved. Finally, further investigation is required to understand the relative importance of subjective probabilities, rather than objective ones examined in this study, about weather events and insurance payout. The attempt to answer these outstanding issues is a promising avenue for future research.

References

Afshar, M. H., Foster, T., Higginbottom, T. P., Parkes, B., Hufkens, K., Mansabdar, S., Ceballos, F., and Kramer, B. (2021). Overcoming basis risk in agricultural index insurance using crop simulation modeling and satellite crop phenology. Working paper, IFPRI, Washington, DC.

Ali, W., Abdulai, A., and Mishra, A. K. (2020). Recent advances in the analyses of demand for agricultural insurance in developing and emerging countries. *Annual Review of Resource Economics*, 12:411–430.

Arnott, R. and Stiglitz, J. E. (1991). Moral hazard and nonmarket institutions: Dysfunctional crowding out or peer monitoring? *The American Economic Review*, 91(1):179–190.

Binswanger, H. P. (1980). Attitudes toward risk: Experimental measurement in rural india. *American Journal of Agricultural Economics*, 62(3):395–407.

Cai, J. (2016). The impact of insurance provision on household production and financial decisions. *American Economic Journal: Economic Policy*, 8(2):44–88.

Cai, J., de Janvry, A., and Sadoulet, E. (2020). Subsidy policies and insurance demand. *American Economic Review*, 110(8):2422–2453.

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

Clarke, D. J. (2016). A theory of rational demand for index insurance. *American Economic Journal: Microeconomics*, 8(1):283–306.

Cole, S., Giné, X., Tobacman, J., Topalova, P., Townsend, R., and Vickery, J. (2013). Barriers to household risk management: Evidence from india. *American Economic Journal: Applied Economics*, 5(1):104–135.

Conradt, S., Finger, R., and Spörri, M. (2015). Flexible weather index-based insurance design. *Climate Risk Management*, 10:106–117.

Dercon, S., Hill, R. V., Clarke, D., Outes-Leon, I., and Taffesse, A. S. (2014). Offering rainfall insurance to informal insurance groups: Evidence from a field experiment in ethiopia. *Journal of Development Economics*, 106:132–143.

Gaurav, S., Cole, S., and Tobacman, J. (2011). Marketing complex financial products in emerging markets: Evidence from rainfall insurance in india. *Journal of Marketing Research*, 48(SPL):S150–S162.

Giné, X., Townsend, R., and Vickery, J. (2008). Patterns of rainfall insurance participation in rural india. *The World Bank Economic Review*, 22(3):539–566.

Giné, X. and Yang, D. (2009). Insurance, credit, and technology adoption: Field experimental evidence from malawi. *Journal of Development Economics*, 89(1):1–11.

Hazell, P. B. R. and Hess, U. (2010). Drought insurance for agricultural development and food security in dryland areas. *Food Security*, 2(4):395–405.

Hess, U. and Hazell, P. (2009). Sustainability and scalability of index-based insurance for agriculture and rural livelihoods. Working paper, IFPRI, Washington, DC.

Hill, R. V., Hoddinott, J., and Kumar, N. (2013). Adoption of weather-index insurance: learning from willingness to pay among a panel of households in rural ethiopia. *Agricultural Economics*, 44(4-5):385–398.

Hill, R. V., Kumar, N., Magnan, N., Makhija, S., de Nicola, F., Spielman, D. J., and Ward, P. S. (2019). Ex ante and ex post effects of hybrid index insurance in bangladesh. *Journal of Development Economics*, 136:1–17.

J-PAL (2016). Make it rain. In *J-PAL Bulletin*. Abdul Latif Jameel Poverty Action Lab, Center for Effective Global Action, and Agricultural Technology Adoption Initiative.

Janzen, S., Magnan, N., Mullally, C., Shin, S., Palmer, I. B., Oduol, J., and Hughes, K. (2020). Can experiential games and improved risk coverage raise demand for index insurance? evidence from kenya. *American Journal of Agricultural Economics*, 103(1):338–361.

Janzen, S. A. and Carter, M. R. (2018). After the drought: The impact of microinsurance on consumption smoothing and asset protection. *American Journal of Agricultural Economics*, 101(3):651–671.

Jensen, N. D., Barrett, C. B., and Mude, A. G. (2016). Index insurance quality and basis risk: Evidence from northern kenya. *American Journal of Agricultural Economics*, 98(5):1450–1469.

Jensen, N. D., Mude, A. G., and Barrett, C. B. (2018). How basis risk and spatiotemporal adverse selection influence demand for index insurance: Evidence from northern kenya. *Food Policy*, 74:172–198.

Karlan, D., Osei, R., Osei-Akoto, I., and Udry, C. (2014). Agricultural decisions after relaxing credit and risk constraints. *The Quarterly Journal of Economics*, 129(2):597–652.

Miura, K. and Sakurai, T. (2015). The impact of formal insurance provision on farmer behavior: Evidence from rural zambia. PRIMCED Discussion Paper 67, Institute of Economic Research, Hitotsubashi University, Tokyo, Japan.

Mobarak, A. M. and Rosenzweig, M. (2012). Selling formal insurance to the informally insured. Working Paper 97, Yale University, New Haven, CT.

Takahashi, K., Noritomo, Y., Ikegami, M., and Jensen, N. D. (2020). Understanding pastoralists' dynamic insurance uptake decisions: Evidence from four-year panel data in ethiopia. *Food Policy*, 95:101910.

Yu, J., Vandeveer, M., Volesky, J. D., Harmoney, K., Yu, J., Vandeveer, M., Volesky, J. D., and Harmoney, K. (2019). Estimating the basis risk of rainfall index insurance for pasture, rangeland, and forage. *Journal of Agricultural and Resource Economics*, 44:179–193.

A Appendix

A.1 Comparative statics of insurance demand

The optimal unit of the insurance contract is $\alpha^* = \frac{p(1-q)-r}{q(1-q)}(y - \underline{y})$ when $p(1-q) - r \geq 0$ holds. The partial derivatives of this demand with respect to p and r are:

$$\frac{\partial \alpha^*}{\partial p} = \frac{1}{q}(y - \underline{y}) > 0$$
$$\frac{\partial \alpha^*}{\partial r} = -\frac{1}{q(1-q)}(y - \underline{y}) < 0.$$

By contrast, the probability of insurance payout, q, has an ambiguous effect on the farmer's insurance uptake. This is because a larger q increases not only the additional cost of insurance purchase (i.e., premium) but also the expected pay-off. The partial derivative concerning q is:

$$\frac{\partial \alpha^*}{\partial q} = \frac{r(1-2q) - p(1-q)^2}{q(1-q)^2}(y - \underline{y}).$$

When $q < 1/2$, the marginal effect could be either positive or negative. (p, q, r) all determine probabilities of false–positive, that is $q - p + r$. Given the pay-off structure and the effect of the false–positive state on the expected utility, the marginal impact of q is positive when r is large enough and negative when p is small enough.

$$\frac{\partial \alpha^*}{\partial q} \geq 0 \leftrightarrow r \geq \frac{p(1-q)}{1-2q} \leftrightarrow p \leq \frac{r(1-2q)}{1-q}.$$

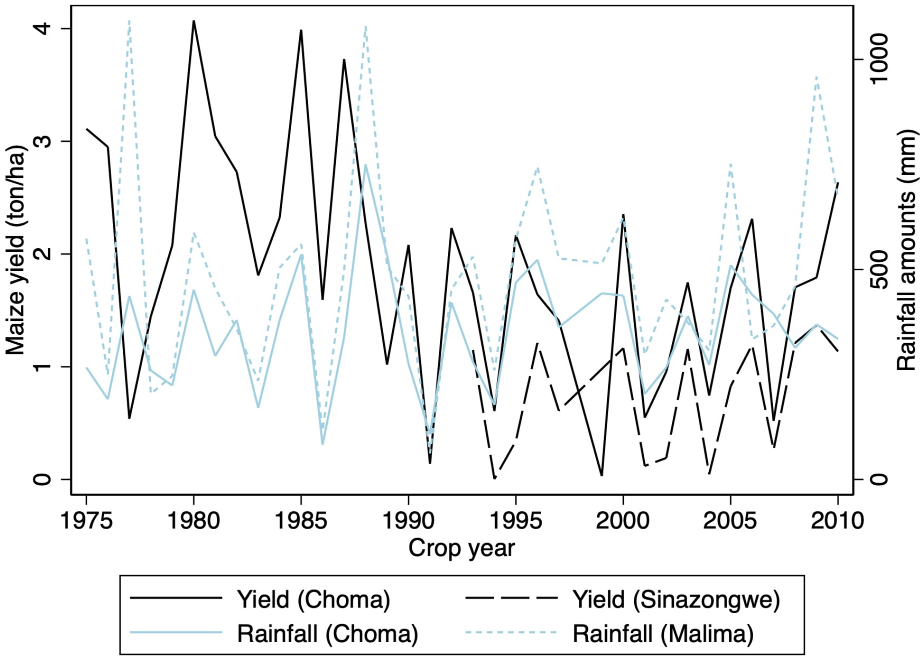


Figure A1: Flowering season's rainfall and maize yield in Choma and Sinazongwe districts, 1975/76–2010/11.

Source: Crop forecast survey data from the Central Statistical Office; Rainfall data of Choma and Malima from the Choma Meteorological Station of Zambia Meteorological Department (Mochipapa) and the Malima irrigation site, respectively.

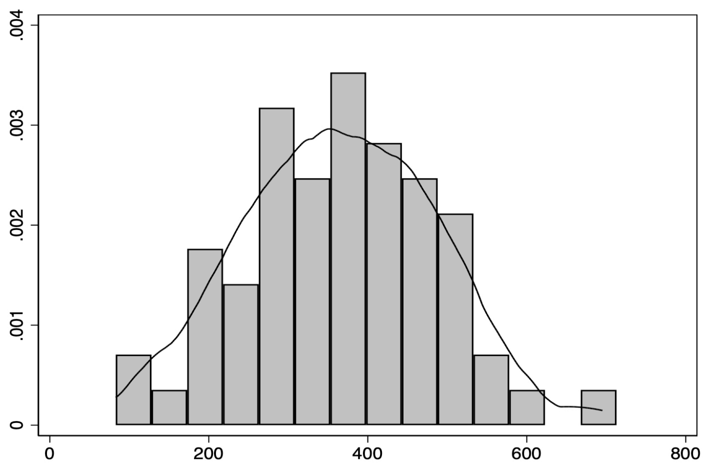


Figure A2: Distribution of rainfall in the flowering season, Choma 1949/50–2011/12

Table A1: Rainfall in Malima and maize yield in Choma and Sinazongwe districts, 1975/76–2010/11

	(1)	(2)	(3)	(4)
Rainfall				
Rainy season (100mm)	0.644***			
	(0.156)			
Rainy season, squared.	-0.030***			
	(0.009)			
Flowering season (100mm)		0.108	0.718***	
		(0.082)	(0.195)	
Flowering season, squared.			-0.055***	
			(0.019)	
Planting season (100mm)		0.035	0.591***	
		(0.073)	(0.151)	
Planting season, squared.			-0.062***	
			(0.015)	
"Drought" in 13/14 contract				-0.699***
				(0.232)
"Flood" in 13/14 contract				-0.933***
				(0.250)
Choma district	0.803***	0.830***	0.846***	0.812***
	(0.202)	(0.210)	(0.186)	(0.158)
Linear time trend	-0.037***	-0.035*	-0.038***	-0.029**
	(0.013)	(0.017)	(0.013)	(0.014)
Constant	-1.349*	1.030**	-1.198**	1.862***
	(0.712)	(0.429)	(0.580)	(0.426)
R-squared	0.527	0.388	0.587	0.454
Observations	51	51	51	51

Notes: The data sources were crop forecast survey data from the Central Statistical Office and rainfall data from the Malima irrigation site. The dependent variable is maize yield (mean = 1.52 tonnes/ha, Std.Dev. = 1.04). The sample covered the period between 1975/76 and 2010/11 for Choma district and between 1993/94 and 2010/11 for Sinazongwe district, with missing observations. "Drought" is a dummy variable that takes the value of 1 if the total rainfall amount during November and December (planting season) was below 214 mm, and 0 otherwise. "Flood" is a dummy variable that takes the value of 1 if the total rainfall amount during November and December (planting season) was above 800 mm, and 0 otherwise. In pooling maize yield data from the Choma and Sinazongwe districts, OLS was used for the estimations. Heteroskedasticity-robust standard errors in parentheses. $*p < 0.1$, $** \ p < 0.05$, $*** \ p < 0.01$.

Table A2: Results of the normality test for rainfall in the flowering season at weather observation points

Test	Statistics	p–value
Choma meteorological station		
Shapiro–Wilk test	$W = 0.986$	$p = 0.688$
Skewness and kurtosis tests	$\chi^2 = 2.320$	$p = 0.313$
Kolmogorov–Smirnov test		
Normal distribution	$D = 0.055$	$p = 0.681$
Log-normal distribution	$D = 0.997$	$p = 0.000$
Observation	64	
Malima irrigation site		
Shapiro–Wilk test	$W = 0.973$	$p = 0.474$
Skewness and kurtosis tests	$\chi^2 = 2.250$	$p = 0.324$
Kolmogorov–Smirnov test		
Normal distribution	$D = 0.131$	$p = 0.246$
Log-normal distribution	$D = 0.971$	$p = 0.000$
Observation	41	

Notes: The table reports results of statistical tests for the normality of total rainfall in January and February for both the Choma weather station and the Malima irrigation site. For the Choma weather station, the rainfall data period used for the tests is between 1949/50 and 2011/12. For the Malima irrigation site, the rainfall data period used for the tests is between 1972/73 and 2012/13. Kolmogorov–Smirnov test compares the observed distribution and the normal (log-normal) distribution using the sample mean and standard deviation.

Table A3: Purchased units of insurance contracts, 2011/12–2013/14

	2011/12		2012/13		2013/14	
	Mean	Std.Dev.	Mean	Std.Dev.	Mean	Std.Dev.
Old sample	2.929	(2.224)	2.614	(1.956)	2.405	(1.231)
Site A	1.857	(1.460)	1.933	(1.032)	2.200	(0.775)
Site B	3.615	(2.219)	3.077	(1.706)	2.231	(1.013)
Site C	3.333	(2.554)	2.875	(2.630)	2.786	(1.717)
New sample	2.589	(2.357)	2.259	(1.884)	2.182	(1.272)
Combined	2.735	(2.295)	2.356	(1.904)	2.243	(1.260)
Total observations	100		160		152	

Table A4: Distribution of purchased units of insurance contracts, 2011/12–2013/14

Units	2011/12	2012/13	2013/14
0	3	1	0
1	6	11	7
2	18	18	20
3	1	2	10
4	10	8	4
6	1	2	0
8	1	1	1
10	2	1	0
Total	42	44	42

Table A5: Test of the model predictions with a restricted sample

	(1)	(2)	(3)	(4)	(5)	(6)
r × Loss	-25.221	-32.843**	-27.947	-34.281**	-25.904	-30.882**
	(17.505)	(14.929)	(17.100)	(16.248)	(17.763)	(14.682)
p × Loss	22.316*	50.080***	23.499**	47.877***	21.978**	60.082***
	(11.403)	(14.776)	(9.951)	(16.775)	(10.156)	(15.974)
r		56.014		51.621		102.063***
		(42.651)		(32.164)		(33.629)
p		-75.406**		-68.810**		-128.611***
		(36.419)		(29.687)		(40.199)
Loss		-3.190		-2.803		-4.515**
		(2.234)		(2.283)		(2.120)
RA–Extreme/Severe			-2.120***	-1.989**	-1.983***	-1.664***
			(0.767)	(0.776)	(0.652)	(0.587)
RA–Intermediate			0.030	0.259	0.412	0.874
			(0.706)	(0.748)	(0.626)	(0.632)
RA–Slight			-0.325	-0.213	-0.169	-0.053
			(0.716)	(0.761)	(0.722)	(0.685)
RA–Neutral			-0.752	-0.591	-0.475	-0.319
			(0.731)	(0.795)	(0.676)	(0.606)
Cash reward			0.014	0.011	-0.001	-0.026
			(0.036)	(0.033)	(0.041)	(0.039)
Year = 2012	-1.112*	-0.385	-1.808**	-1.115	-1.740**	0.003
	(0.599)	(0.974)	(0.675)	(0.892)	(0.680)	(0.996)
Year = 2013	-0.137	-1.583	-0.703	-2.018**	-0.665	-3.375***
	(0.633)	(1.248)	(0.644)	(0.943)	(0.660)	(0.986)
Constant	2.198***	8.139**	3.422***	8.668***	3.791***	10.854***
	(0.502)	(3.064)	(0.907)	(2.778)	(0.840)	(2.718)
Household characteristics	No	No	No	No	Yes	Yes
R-squared	0.209	0.256	0.408	0.446	0.450	0.524
Observations	82	82	82	82	82	82

Notes: The estimation sample is restricted to observations in sites B and C. The dependent variable is the number of insurance contracts that a household purchased. RA denotes risk-aversion class and RA–Moderate is used as a reference. Heteroskedasticity-robust standard errors in parentheses. Household characteristics include the household head's sex, age, and years of schooling. $*p < 0.1$, $**\ p < 0.05$, $***\ p < 0.01$.

Table A6: Other potential constraints to insurance demand

	(1)	(2)	(3)	(4)	(5)
Literacy	0.323**	0.302**	0.329	-0.430	-0.382
	(0.137)	(0.136)	(0.230)	(0.777)	(0.672)
Math	0.408***	0.366***	0.347	0.478	0.517
	(0.110)	(0.109)	(0.235)	(0.736)	(0.689)
Risk-sharing	0.095	0.071	-0.013	-0.218	-0.380
	(0.103)	(0.097)	(0.070)	(0.288)	(0.265)
Drought	0.062	-0.055	0.490	1.009	0.406
	(0.367)	(0.357)	(0.915)	(1.884)	(1.413)
r				-36.391*	-26.659
				(19.130)	(20.521)
Literacy × r				6.526	4.868
				(6.490)	(5.611)
Math × r				-1.126	-4.703
				(5.339)	(5.570)
Risk-sharing × r				1.792	2.397
				(2.154)	(1.985)
Drought × r				-3.987	1.234
				(12.257)	(11.637)
r × Loss					-66.941***
					(19.568)
p × Loss					64.119***
					(15.282)
Cash reward		0.053**			0.088**
		(0.021)			(0.042)
Old sample	0.188	0.163			
	(0.339)	(0.333)			
Year = 2012	-0.755	-0.856*	-0.736	-0.853	-2.207***
	(0.484)	(0.493)	(0.716)	(0.716)	(0.613)
Constant	0.658	0.487	1.086	5.345**	5.417*
	(0.558)	(0.753)	(0.647)	(2.328)	(2.865)
Sample	Combined	Combined	Old	Old	Old
Risk-aversion class	No	Yes	No	No	Yes
R squared	0.104	0.144	0.103	0.155	0.541
Observations	260	260	86	86	85

Notes: The estimation sample is limited to the 2011/12 and 2012/13 crop years. OLS was used for the estimation. Heteroskedasticity-robust standard errors in parentheses. *$p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.
Variable: Literacy–a total score of the five questions assessing the comprehension of the insurance product (mean=3.77/5). Math–a total score of three basic arithmetic questions (mean=1.97/3). Risk-sharing–the number of people who can be called in times of need (mean=3.40). Drought–a dummy variable for the experience of drought in the previous crop year (mean=0.10 in 2011/12 and 0.82 in 2012/13).

Table A7: Effect of the basis risk on insurance demand: Fixed effect estimates–bootstrap standard errors

	(1)	(2)	(3)	(4)	(5)
r × Loss	-14.354*	-14.054*	-16.217*	-16.107*	
	(7.403)	(8.206)	(8.405)	(8.967)	
r		-4.801		-1.844	-19.765
		(19.851)		(18.520)	(19.464)
RA – Extreme/Severe			-0.582	-0.592	-0.518
			(0.411)	(0.430)	(0.446)
RA – Intermediate			0.519	0.511	0.349
			(0.553)	(0.559)	(0.521)
RA – Slight			0.900	0.891	0.871
			(0.567)	(0.584)	(0.552)
RA – Neutral			0.288	0.280	0.151
			(0.569)	(0.580)	(0.549)
Cash reward			0.026	0.026	0.023
			(0.029)	(0.029)	(0.035)
Year = 2012	-1.188**	-1.257**	-1.550**	-1.577**	-0.894
	(0.565)	(0.626)	(0.662)	(0.700)	(0.553)
Year = 2013	-0.231	-0.148	-0.401	-0.369	-0.422
	(0.357)	(0.422)	(0.417)	(0.453)	(0.479)
Constant	4.132***	4.493***	4.149***	4.299**	4.442**
	(0.727)	(1.589)	(1.001)	(1.684)	(1.901)
R squared	0.678	0.678	0.739	0.739	0.680
Observations	126	126	126	126	126

Notes: The dependent variable is the number of insurance contracts that a household purchased. RA denotes risk-aversion class and RA–Moderate is used as a reference. Robust standard errors computed using 1000 bootstrap replications in parentheses. $*p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.