Paying Premiums with the Insurer's Money: How Loss Aversion Drives Dynamic Insurance Decisions

Daniel Stein (London School of Economics)*

Job Market Paper, October 2011

Abstract

This paper anlyzes the dynamic nature of rainfall insurance purchasing decisions, specifically looking at whether and why receiving an insurance payout induces a greater chance of purchasing insurance again the next year. Using customer data from the Indian micro-finance institution BASIX, I find that receiving an insurance payout is associated with a 9-22% increased probability of purchasing insurance the following year. These empirical results are consistent with the predictions of a loss aversion model where receiving insurance payouts shifts a customer's reference point, resulting in higher risk aversion and therefore greater demand for insurance. I do not find significant support for other potential mechanisms, such as previous weather directly affecting insurance decisions or payouts increasing trust in the insurance companies. Overall, low repurchasing rates even after payouts suggest that current rainfall index insurance products are likely to continue struggling in their current form.

1 Introduction

Roughly 60% of India's population is employed in agriculture, and over 50% of agricultural land is dependent on rainfall to nurture the crops.[1] But the Indian monsoon is notoriously unpredictable, prone to droughts and floods that can have devastating effects on the livelihood of rural Indians. While Townsend (1994) argues that Indian villages do an effective job of providing informal consumption insurance against idiosyncratic shocks, a poor monsoon will hit whole villages and districts at once, likely rendering intra-village transfers ineffective. Beginning in the early 2000s, rainfall index insurance

*This work would not have been possible without the assistance of the insurance staff at BASIX, especially Sridhar Reddy, for help in obtaining and understanding their customer database. Also thanks to Tim Besley, Greg Fischer, and Gani Aldashev for helping direct the project and commenting on earlier drafts. Thanks to participants at the Munich Re 6th Annual Microinsurance Conference, the Midwestern International Economic Development Conference, seminars at the London School of Economics, University of Namur, and KU Leuven for helpful comments.

[1] CIA World Factbook: India; Indiastat.com

was introduced in India as a potentially important tool to help poor farmers deal with rainfall risk (Hess, 2004), but has struggled to reach a critical mass of customers, especially when unsubsidized.[2]

Unlike physical goods (or even credit), it is difficult for customers to evaluate the benefits of insurance since its main benfits only occur when a payment is received. If customers are unfamiliar with how insurance works, they may be influenced by their recent experiences with insurance and also by the experiences of their friends and neighbors. Providing evidence from the developed world, Kunreuther et al. (1985) observe that purchases of flood and earthquake insurance in the US are greatly influenced by recent experiences with disasters and insurance payouts, and also argue that peoples' insurance decisions are influenced by their friends and neighbors' experiences with insurance. Reacting to low rates of rainfall insurance uptake in Andhra Pradesh, India, Giné et al. (2008) suggest that "over time, lessons learned by insurance 'early adopters' will filter through to other households, generating higher penetration rates among poor households."

This paper seeks to understand how previous insurance payouts can affect future insurance purchasing decisions, and what mechanisms can explain this behavior. Using data on three years of insurance purchasers from the Indian micro-finance institution BASIX, I find that customers who received an insurance payout are 9-22% more likely to repurchase in the following year as compared to customers who did not receive any insurance payments. Customers who received larger insurance payouts are more likely to repurchase than those who received small payouts. However, the paper finds no evidence of positive spillover effects to other people in the village even at large levels of payouts, casting doubt on the hypothesis that witnessing insurance payouts will spur new buyers.

I introduce a model in which customers exhibit loss aversion and repurchases of insurance are driven by the psychological effects of a reference point shift. Purchasers of insurance who receive insurance payouts will be more likely to purchase insurance in the future because their previous insurance "profits" make future premiums seem like less of a loss. Following the logic of Thaler and Johnson (1990), I propose that customers who receive an insurance payout will not regard future premium payouts as a true loss, which can be modeled as a shift of the reference point that increases risk aversion. I define the conditions under which risk aversion will increase after a payout, and argue that these conditions are likely to hold.

I also explore some alternative hypotheses as to why receiving payouts could increase insurance demand the following year. First, it is possible that weather shocks themselves could have an effect on insurance demand, such as Kunreuther et al. (1985) observe in the US. This could happen because weather shocks change customers' beliefs about future shocks, change their wealth, or simply increase the salience of shocks. I look for direct effects of weather by testing how rainfall in the year before insurance was introduced affected insurance purchases, and find evidence that previous rainfall shocks

[2] There have been some areas with a large number of insured, but this is generally due to heavy government subsidies or requiring insurance to obtain government loans. Sending a strong signal about official uncertainty about the prospects of rainfall index insurance, the government of India decided to delay expansion of its rainfall index insurance pilot (WBCIS) to all of India, and in 2010 piloted a modified version of its area-yield index product (NAIS).

tend to *decrease* insurance purchasing. This provides evidence against the argument that it is weather shocks as opposed to insurance payouts that are driving insurance purchases.

Next I test whether receiving insurance payouts could induce trust in the insurance company or learning about the insurance product. Bryan (2010) suggests that index insurance take-up is low due to ambiguity aversion, which should decrease as customers learn more about the product. We assume that if trust and learning are driving purchases, we should be able to witness spillover effects on other people in the village. This is because people in a village who witnessed payouts but did not receive them should also have been able to learn about insurance and gain trust in the insurance company. I do not find convincing evidence that these spillover effects are present, and argue that this is evidence that insurance repurchasing is not being driven by trust or learning. Overall, the empirical analysis suggests that it is the physical reception of payout money that drives future purchases, which is consistent with the proposed model of loss aversion.

This paper is related to a few separate lines of research. First, it contributes to a growing list of empirical studies that attempt to determine demand for weather index insurance (Hill and Robles, 2010; Cole et al., 2010; Giné et al., 2008). One overarching conclusion from previous studies on index insurance is that demand for index products is low when provided at market rates[3], and only increases when prices are slashed significantly. However, most of these studies look at insurance as a static purchasing decision, seeing what factors lead people to become first time customers.

One exception is Hill and Robles (2010), who provide rainfall insurance for free as part of an experimental game in Ethiopia, and then return the next year to sell the same insurance. Despite the fact that two-thirds of the people who were granted insurance during the experiment received payouts, this group had low take-up rate of 11% the next season, which was a lower rate than those who had not participated in the experiment.

Our study differs from these in that it uses a much larger dataset, allowing comparisons across weather stations to identify the dynamic effects, which I argue allows a clearer identification of the effects of payouts. Additionally, this paper studies a real world insurance implementation at market rates, making its results potentially more relevant for policy making.

Another related work is an observational study on mutual insurance among fisherman in the Ivory Coast by Platteau (1997). Platteau observes malfunctioning mutual insurance cooperatives and theorizes that they are failing because members view insurance as a system of balanced reciprocity, meaning that they expect to break even over the lifetime of the scheme. When members have not received the services (in this case sea rescue) of the mutual in a long time, they start to view the insurance as a bad deal and ask for their contributions back. This paper's loss aversion model provides a different theoretical framework that can also generate predictions of a desire for balanced reciprocity among insurance customers.

This work also contributes to the literature on choice under uncertainly by providing a theoretical

[3]Market rates tend to be around 2-6 times actuarially fair rates.

framework to understand how shifting reference points can help explain dynamic insurance choices. The basic ideas behind the model are not new: in their original paper on prospect theory Kahneman and Tversky (1979) note that if customers "expect" to purchase insurance such that their payment of the insurance premium is not counted as a loss, this makes insurance more attractive. Similarly, Köszegi and Rabin (2007) argue that if a risk and the ability to insure it are "anticipated," then payment of the insurance premium should be incorporated into the reference point and therefore not counted as a loss.

But neither of these papers are specific about what may cause a customer to move from one reference point to another. Kaheneman and Tverky muse that someone may expect to purchase insurance (and therefore have a different reference point) "perhaps because he has owned it in the past or because his friends do." This paper builds on these ideas but adds the crucial assumption that receiving insurance payouts is the key event that can trigger a change in the reference point. One way to justify this assumption is to invoke the concept of the "lagged status quo," which means that after experiencing a gain (or loss), the new reference point may not immediately update to the current level of wealth. If the reference point lags after an insurance payout, loss aversion dictates that people will become more risk averse and therefore more likely to purchase insurance in the next period.

Behavior consistent with the lagged status quo is demonstrated by Thaler and Johnson (1990), who show that gamblers who had experienced recent gains were less sensitive to subsequent losses. (This is the "gambling with house money" effect.) Their framework offers the equivalent of a reference point shift such that after winning money subjects do not psychologically consider subsequent negative outcomes as true "losses." When the same logic is applied to insurance purchases, it means that people who receive an insurance payout will not regard subsequent premium payouts as a loss, making them more risk averse and therefore more likely to purchase insurance.[4] This is the mechanism utilized in the theoretical framework of this paper.

Another relevant work is Köszegi and Rabin (2007), who adapt their theory of stochastic reference points (Köszegi and Rabin, 2006) to insurance decisions and explicitly outline how different reference points will affect insurance decisions. In Köszegi and Rabin's model the "correct" reference point to consider when analyzing insurance decisions depends on whether the gamble is anticipated, and how long a time there will be between when the gamble is chosen and the payouts are realized. Their key insight as relates to this study is that if insurance purchases are anticipated and made far in advance of their outcome, they will not be counted as a loss as the reference point adjusts to account for the insurance premium. If the transition from a "surprise" gamble to an "anticipated" expense happens after customers receive insurance payouts, Köszegi and Rabin's model would generate predictions similar to this paper. However, this is difficult to justify because insurance decisions are made over

[4]The subjects in Thaler and Johnson's experiments play gambles involving gains, so shifting the reference point makes them more risk loving, as potential losses are now less painful while gains are relatively unchanged. However, when the same principle is applied to insurance, a gamble which involves only losses, the same reference shift will make customers more risk averse.

the same time frame each year. The model presented here instead relies on a simpler mechanism, assuming that reference points are fixed within a decision period but can shift based on the outcome of that period.

The paper will proceed as follows. Section 2 outlines the model and determine the conditions under which an insurance payout will lead to greater insurance purchasing. The model predicts that insurance customers who receive a payout will be more likely to purchase insurance the following year. Section 3 explains the insurance policies and data that will be studied in the empirical section. Section 4 provides the main empirical evidence, which confirms the main predictions of the theory. Section 5 looks at alternative explanations for the increased insurance purchases, such as the possibility that they are driven by increased trust or pure weather effects. I look for evidence that these mechanisms are driving insurance decisions and fail to find them. Section 6 concludes and offers policy recommendations based on the results.

2 Theory

In this section I introduce a theoretical framework that seeks to explain how experience with insurance could affect the decision to purchase insurance during the following season. I present a model where agents exhibit loss aversion, but their reference point can shift based on previous experiences with insurance. The key component of the model is the reference point shift, which can alter the risk aversion of the agents and therefore change their demand for insurance. As the empirics of this paper show that receiving a payout correlates with increased insurance purchases, I focus on payout reception as the key moment when reference points are likely to shift. After setting up the basic framework of the model, I then determine the restrictions on the evolution of the reference point that will yield the aforementioned empirical prediction.

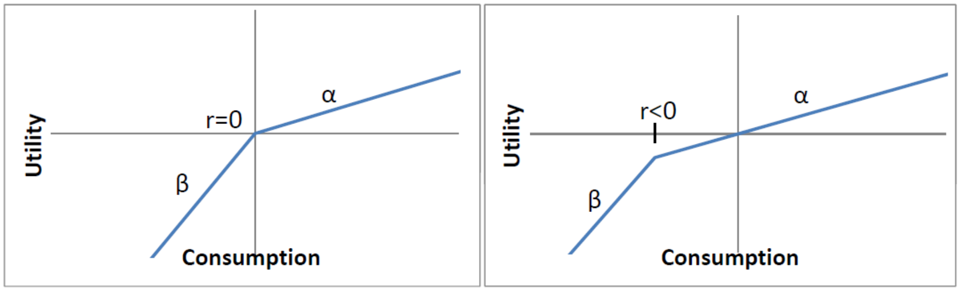
Subjects have a piecewise-linear utility function that exhibits loss aversion around a reference point. Given a reference point r, utility is:

$$u(c,r) = \begin{cases} \alpha c & if \quad c > r \\ \beta(c-r) + \alpha r & if \quad c < r \end{cases} \tag{1}$$

Equation 1 defines the utility funtion for $r < 0$, as this is where all the interesting dynamics of the model take place. Figure 1 shows the utility function for two values of r.

The model lasts two periods, and in each period there are two possible states of the world $S = \{0, 1\}$. If $S = 1$, which happens with probability p, agents suffer a consumption shock of $-X$. If $S = 0$ there is no shock. Agents also have the opportunity to purchase insurance $I = \{1, 0\}$ against the shock. If

Figure 1: Utility Functions



The function on the left has reference point $r = 0$, and is risk neutral for losses. The figure on the right has $r < 0$, making the subject risk averse for losses.

an agent purchases insurance he is completely protected from the consumption shock should $S = 1$ occur. Insurance costs a constant multiple $(1 + \lambda)$ of the expected payout of insurance, resulting in an insurance premium of $(1 + \lambda)pX$. At the end of period 1 the agent's reference point can move, as defined by the function $r_2 = f(r_1, I_1, S_1)$. This means that the reference point has a chance to shift from the first to the second period based on a customer's experience with insurance (determined by I_1, S_1) in the first period. The form of the function $f(r, I, S)$ will determine the interesting dynamics of the model, as it is the change in reference point that will generate changes in the insurance decision.

Timing is as follows:

1. Agent starts with reference point r_1
2. Agent chooses insurance decision $I_1 = \{0, 1\}$.
3. State of the world $S_1 = \{0, 1\}$ is realized. Agent receives period 1 utility.
4. Agent's reference point moves to $r_2 = f(r_1, I_1, S_1)$
5. Agent chooses $I_2 = \{0, 1\}$
6. State of the world $S_2 = \{0, 1\}$ is realized, and agent receives period 2 utility.

There is no discounting. Each period, the agent decides whether or not to purchase insurance by considering his expected utility with or without insurance. I assume that $0 \geq r \geq -X$, which restricts analysis to the case where some loss aversion is present over the possible range of consumption.[5] I also assume that agents are naive, which means that they don't take possible shifts in their reference point into account when making their first period insurance decision.[6]

[5] $r = 0$ and $r = -X$ give risk neutral preferences, but I include these endpoints as limits of the interesting range of r.

[6] A sophisticated agent would make different choices in period 1 since they would anticipate how those choices would change their reference point in period 2. However, the difference in choices in the second period between an insurance customer who received a payout in period 1 versus a customer who did not receive a payout will not change if agents are

Applying the utility function defined in Equation 1, expected utility U in each period is defined in Equation 2.

$$\begin{cases} U(r,I) = \beta I(((-(1+\lambda)pX) - r) + \alpha r) + p(1-I)(\beta(-X-r) + \alpha r) & if \quad -(1+\lambda)pX < r \\ U(r,I) = \alpha I(-(1+\lambda)pX) + p(1-I)(\beta(-X-r) + \alpha r) & if \quad -(1+\lambda)pX > r \end{cases} \tag{2}$$

In this simple decision model, an agent would choose to purchase insurance if expected utility from purchasing insurance was greater than expected utility from forgoing insurance. However, in a richer model there may be other factors that would influence a potential buyer's purchasing decision. I therefore analyze the benefit of insurance, defined as the difference in expected utility from purchasing insurance versus forgoing insurance. I assume that if people have greater benefits from purchasing insurance, they will be more likely to purchase.

$$B(r) \equiv U(r,1) - U(r,0) \tag{3}$$

One can interpret the benefits of insurance $B(r)$ to be a measure of risk aversion, as higher benefits from insurance imply higher risk aversion.[7]

The Period 2 reference point r_2 is determined by the initial reference point, insurance decision, and state of the world in period 1 according to $r_2 = f(r_1, I_1, S_1)$. We can use the function f to connect benefits of insurance in Period 2 to insurance experiences in Period 1. People who received a payout in period 1 will be more likely to re-purchase insurance than insurance customers who did not receive a payout if

$$B_2(f(r_1,1,1)) > B_2(f(r_1,1,0)) \tag{4}$$

The key is to discover the conditions on $f(r,I,S)$ such that Inequality 4 holds. Due to the piecewise definition of the utility function, the complete set of conditions that guarantee Inequality 4 holds are somewhat complex and are explained in detail in the Appendix. In order to illustrate the basic mechanism we can examine the specific and highly plausible case where $r_1 = 0$ and $f(r_1,1,0) = r_1 = 0$. The assumption of $r_1 = 0$ can be interpreted as the agent having no previous experience with insurance, and therefore exhibiting standard loss aversion with the reference point being his initial level of consumption.[8] $f(r_1,1,0) = r_1$ means that insurance purchasers who do not receive an insurance

sophisticated. As this is the main result we care about, considering only naive agents lends considerable mathematical simplicity to the problem.

[7] $B(r)$ as defined here will be monotonically related to the risk premium as defined in Pratt (1964).

[8] If one follows Köszegi and Rabin (2006), the initial reference point would be the expectation of consumption, making $r_1 = -pX$. The main intuition that follows is still valid, and the solution to Inequality 4 for a general $r_1$is given in Proposition 2 in the Appendix. One point to notice is that the maximum benefit of insurance is obtained when $r = -\lambda pX$ (Proposition 7.1 in the Appendix). Therefore, if $\lambda \leq 1$ and the intial reference point is the expectation of consumption,

payout do not experience a shift in their reference point.

Proposition 1. In the case of $r_1 = 0$, $f(r_1, 1, 0) = r_1$, $B_2(f(0,1,1)) > B_2(f(0,1,0))$ iff $0 > f(0,1,1) > -(1+\lambda)X$.

In words, this proposition states that if an insurance payout causes a decrease in the reference point (up to $-(1+\lambda)X$), this increases the agent's risk aversion, and therefore his benefits from purchasing insurance in period 2. Intuitively, a decrease in the reference point corresponds to a lagged status quo, and makes premium payments in the next period not seem like a real loss.

Proof. I start by looking at people who purchased insurance in period 1 but did not receive a payout. Since we have assumed that for this group the reference point remains at zero, the benefits of insurance in period two are straightforward to calculate from Equations 2 and 3.

$$B_2(0,1,0) = \lambda\beta pX \tag{5}$$

For those who did receive a payout, the reference point will move to $r_2 = f(0,1,1)$. Due to the piecewise nature of the utility function, the benefits are also going to depend on whether or not the reference point remains above the insurance premium or moves below it. From Equations 2 and 3 we see:

$$\begin{cases} B_2(0,1,1) = \lambda\beta pX + (1-p)(\alpha-\beta)f(0,1,1) & if \quad 0 > f(0,1,1) > -(1+\lambda)pX \\ B_2(0,1,1) = -(1+\lambda)\alpha pX + p\beta X - p(\alpha-\beta)f(0,1,1) & if \quad f(0,1,1) < -(1+\lambda)pX \end{cases} \tag{6}$$

In the case that $0 > f(0,1,1) > -(1+\lambda)pX$, $B_2(0,1,1) > B_2(0,1,0)$ iff $f(0,1,1) < 0$. When $f(0,1,1) < -(1+\lambda)pX$, we can prove the inequality by combining Equations 6 and 7.

$$B_2(0,1,1) - B_2(0,1,0) = p(\beta-\alpha)(f(0,1,1) + (1+\lambda)X) \qquad if \qquad f(0,1,1) < -(1+\lambda)pX \tag{7}$$

So for this case $B_2(0,1,1) > B_2(0,1,0)$ iff $f(0,1,1) > -(1+\lambda)X$. Combining these two cases proves Proposition 1. □

In words, this means that buyers who receive insurance payouts in period 1 will be more likely to purchase insurance in period 2 if their reference point decreases as a result of receiving a payout. This

the reference point yielding maximum benefit from insurance lies above the starting reference point, and the dynamics of the model change considerably. In the real world this is an unlikely scenario as market prices for insurance are generally above their expected payout, making $\lambda > 1$. In our empirical sample of BASIX insurance products, λ ranges from 1.1 to 15.7, with an average of 5.8.

basic result should be apparent from analyzing Figure 1. In the left panel of Figure 1, the reference point is at zero and the subject is risk neutral for losses in relation to the reference point. In the right panel, the reference point has moved to below zero and the customer is now risk averse, therefore showing increased demand for insurance.

The second condition $f(0,1,1) > -(1+\lambda)X$ will only bind if $\lambda < 0$, as we have already assumed that $r > -X$. This condition comes about from the fact that when $\lambda < 0$ there is potentially a benefit from purchasing insurance even for risk lovers, as the expected return from purchasing insurance is higher than the expected returns without. I'll call this the expectation benefit. The expectation benefit is greater if an agent's utility function is steeper between the expected returns with and without insurance. When the reference point is below the price of insurance, the utility function flattens in this range, decreasing the expectation benefit. As the reference decreases from $-(1+\lambda)pX$ towards $-X$, risk aversion decreases until the utility gains from insurance approach the utility losses from the reduced expectation benefit. These forces are equal when $f(0,1,1) = -(1+\lambda)X$.

From Proposition 1, we see that if the reference point moves such that it is below the level of current wealth, customers become more risk averse and therefore more likely to purchase insurance. Such a shift is likely to occur after an insurance payout assuming that insurance customers behave in a manner consistent with the participants in Thaler and Johnson's (1990) lab experiments. In these experiments people were more likely to gamble with money they had recently won as they did not regard losing this money as a true loss, which is consistent with having a reference point below the level of current wealth. I hypothesize that people who receive an insurance payout regard this payout akin to a gambling victory, and therefore this payout shifts their reference point below the level of current wealth. According to Proposition 1, this shift makes recipients of insurance payouts more likely to purchase insurance than other customers who did not receive payouts.

While this model is similar to models of the lagged status quo (Thaler and Johnson, 1990; Gomes, 2005), it is formulated somewhat differently to allow for greater flexibility in the movement of the reference point. In standard models of the lagged status quo, the reference point does not move after a gamble even though wealth has changed, causing a wedge between current wealth and the reference point. In this model, at the beginning of each period utility is normalized to zero, while instead the reference point is allowed to shift. Despite this change in modeling form, this model also creates a wedge between current wealth and the reference point, allowing it to give similar predictions as a lagged status quo model. For instance, if the reference point shifts down after experiencing a gain then the reference point is below the current level of wealth, just as it would be in a lagged status quo model. Allowing the reference point to shift arbitrarily, however, gives this model the ability to both reproduce the results of lagged status quo models while also allowing insights gained from other possible reference point changes.

This theory provides a framework that we can use to analyze how insurance experiences can change future insurance decisions. Specifically, it offers conditions under which customers who receive payouts would be more likely to purchase insurance in the following period. Assuming these conditions hold, this model predicts that people who receive a payout will be more likely to purchase insurance in the following year. I turn next to the empirical section to show that the behavior of rainfall insurance purchasers in India is indeed consistent with this prediction. I then show that alternative mechanisms which would fall outside of this theory, such as purchases being driven by increased trust in the insurance company or direct effects of weather, are not supported by the data.

3 Index Insurance and Customer Data

3.1 Context: BASIX Policies

In this analysis I study monsoon rainfall index insurance policies underwritten by ICICI-LOMBARD and sold by BASIX, a microfinance institution based in Hyderabad. The policies insure against excess or deficit rainfall, and are calculated based on rainfall measured at a stated weather station. By basing payoffs on just rainfall, the policies should have low monitoring and verification costs, and also should be free of adverse selection and moral hazard. These attributes make policies inexpensive to create and administer, which allows them to be sold in small quantities and priced at levels affordable for poor farmers. BASIX's policies are designed to pay out in situations where adverse rainfall would cause a farmer to experience crop loss, and are therefore calibrated to the water needs of local crops.

BASIX policies are divided into three phases, which are meant to roughly capture the three phases of the growing season: planting, budding/flowering, and harvesting. If cumulative rainfall is too low or high in any of these phases, the crop's output is potentially damaged and the farmer could suffer a loss. The policies are designed to start when farmers first start planting, which depends itself on rainfall. Therefore, the policies have a dynamic start date which means that Phase 1 begins on the day that cumulative rainfall since June 1 reaches 50mm or on July 1, whichever comes first. Each phase generally lasts 35-40 days. During this time, rainfall data is collected daily at a designated weather station, and payouts are calculated using the cumulative rainfall over the phase.

A phase of coverage is defined by three parameters: "Strike", "Exit", and "Notional". Deficit policies begin to pay out when the rainfall drops below the level of the Strike, and gives its full payout when it falls below the Exit. In between, it pays the Notional amount of rupees for each millimeter below the Strike.

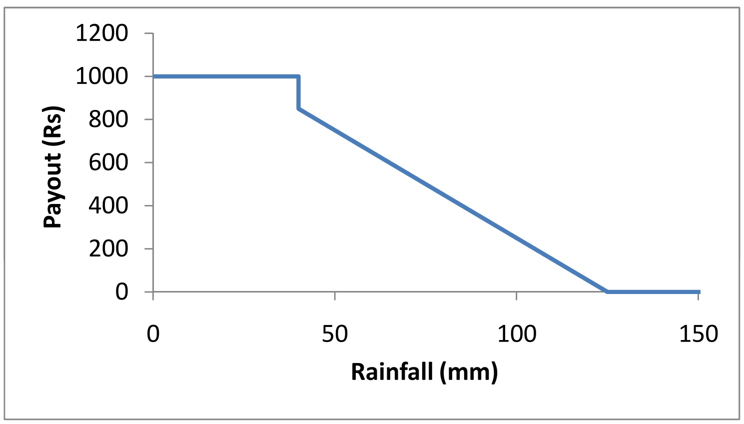
In 2006 and 2007, all rainfall insurance contracts sold by BASIX included three phases, with the first two protecting against deficit rainfall, and the third protecting against excess rainfall. In 2005 the policies all had three phases, but each phase protected only against deficit rainfall. Table 1 presents a sample contract, from Nizamabad district in the state of Andhra Pradesh.

Given the policy parameters we can see how the payouts will evolve according to rainfall. Figure 2 shows the payout schedule for phase II of the above policy. There is no payout when rainfall is above

Table 1: Example Insurance Policy

Phase	I	II	III
Duration (Days)	35	35	35
Type	Deficit	Deficit	Excess
Strike (mm)	135	125	730
Exit (mm)	40	40	820
Notional (Rs/mm)	10	10	10
Policy Limit (Rs)	1000	1000	1000
Premium (Rs)	110	110	90

Figure 2: Example Payout Schedule



the strike, which is 125mm. Then as rainfall decreases the payout increases linearly until rainfall reaches the exit of 40mm, then jumps to the policy limit of Rs 1000 once rainfall falls below 40mm.

BASIX insurance policies are sold in April and May, which are the months that precede the monsoon in India. Insurance policies cover only one season, so customers must purchase insurance again if they want coverage for the following year.

Table 2 presents summary statistics for the insurance policies studied.

3.2 Data

The data set consists of the entire set of BASIX's purchasers of rainfall index insurance from 2005-2007, which covers six states.[9] Though it ran small pilots in 2003 and 2004, BASIX began to mass-market

[9]The states are, in descending order of number of buyers: Andhra Pradesh, Maharashtra, Jharkand, Karnataka, Madhya Pradesh, Orissa.

Table 2: Policy Summary Statistics

Year	2005	2006	2007
Number of Policies	34	42	28
Average Premium For Three Phases (Rs)	283	295	287
Expected Payout (Using rainfall from 1961-2004)	76	73	80
Mean Ratio of Premium to Expected Payout	5.76	5.94	5.6
Mean Percentage of Times Policy Paid out From 1961-2004	10.2	6.8	7

rainfall insurance starting in 2005. The data contains limited personal information about each customer including their location, how many policies they purchased, and what payouts they received during that season. The BASIX data covers 42 weather stations, and includes a total of 19,882 customers from 2005-2007.[10] After numerous rainfall shocks in 2006, BASIX realized that many customers who had purchased only a small amount of insurance were disappointed that they received small payouts. In response, BASIX instituted a rule in 2007 that required all customers to purchase insurance coverage with a maximum payout of at least Rs 3000. This was meant to encourage people to buy a level of coverage that would actually provide meaningful payouts in the event of a shock, but resulted in a sharp decrease in the number of customers in 2007. A summary of characteristics of BASIX customers is given in Table 3.

For rainfall data, I use a historical daily grid of rainfall, which is interpolated based on readings from thousands of rainfall stations throughout India. This data is provided by the Asian Precipitation Highly Resolved Observational Data Integration Towards Evaluation of water resources.[11] This data set has daily readings of rainfall from 1961-2004, at a precision of .25°[12]. For each .25° x .25° block, the data contains the amount of rainfall in millimeters and the number of stations within the grid that contributed to the data. This data is used to evaluate how the insurance policies would have paid out historically, which can be used as a proxy for past rainfall shocks.

The initial challenge in processing the BASIX administrative data was to turn it into a panel. Although BASIX had three years of data, there was no way to identify unique individuals that purchased in multiple years. In order to solve this problem I matched names from year to year, taking into account the customer name, father/husband name, location, and age in order to identify which households were the same.[13]

[10]Note that BASIX also sold many policies in the district of Deogarh in Jarkhand, and those buyers are omitted from this analysis. The reason for this is that the policy for Deogarh is heavily subsidized, resulting in a policy that is completely different from all the others. For instance, the Deogarh policy for 2005 has an expected payout of Rs 1140 compared to an average of Rs 149, although the policy does not cost more than average. Because of its incredibly generous terms, the Deogarh policy has huge payouts for all years of the study, and therefore does not seem to be 'normal' enough to warrant inclusion in the main dataset. All the analysis below is performed excluding all buyers in Deogarh, though most results do not change substantially when it is included.

[11]APHRODITE's water resources project; http://www.chikyu.ac.jp/precip.

[12].25° Latitude equals about 27.5km, while .25° longitude varies by latitude. Over the range of latitudes in this survey it equals roughly 26km.

[13]As district, village, and customer names had highly variant spelling, it wasn't possible to match customers through

Table 3: Customer Summary Statistics

Year	2005	2006	2007
Number of Villages	954	1426	432
Number of Weather Stations	34	42	28
Number of Buyers	6428	10077	3377
Average Sum Insured (Rs)	3055	1612	3547
Buyers Receiving Payouts	351	1346	529
Buyers Who bought the Following Year	453	364	

The potential for matching errors causes a serious concern about the validity of the data set. Since a crucial part of my analysis revolves around determining what causes buyers to re-purchase insurance, determining who does so is extremely important. Despite a comprehensive effort which combined automated and manual matching methods, there are certain to be some errors in the dependent variable. While there is no reason to believe that this measurement error is correlated with any independent variables in the regression, since the dependent variable in some regressions is a dummy variable this can lead to downward bias on the estimated coefficients. In Section 4 I will explore the possible consequences of this problem.

A more serious problem with the data set is that it is not possible to observe the level of marketing that each person received, making "marketing intensity" an important omitted variable. When BASIX markets rainfall insurance, it first calls a group meeting in a village, and shows the villagers a video about rainfall insurance (and other BASIX products). It then speaks with visitors and answers questions. The BASIX team then makes a follow-up visit where it goes door to door, trying to sell BASIX products including rainfall insurance. Unfortunately, I have no data on on the specific marketing practices of each village and don't even know for sure in which villages BASIX actively sold rainfall insurance each year. As marketing intensity is potentially correlated with previous insurance outcomes, this may bias the estimates. This needs to be taken into account when performing the analysis and interpreting the results.

4 Results: The Effect of Payouts on Take-up

In this section I address the central question: is receiving an insurance payout correlated with repurchasing insurance the following year? To do this I examine BASIX's customers in 2005 and 2006, and regress repurchasing on payout reception and a year dummy. The basic econometric specification is as follows:

$$y_{i,t+1} = \alpha + \beta_1 P_{i,t} + B_2 D_{2006} + \epsilon_{t,i} \tag{8}$$

the years using automated means.

Table 4: Insurance Repurchasing

	Dependent Variable is Customer Re-Purchasing Insurance			
	(1)	(2)	(3)	(4)
Received Insurance Payout	0.0897***	0.222***	-0.0877***	-0.195***
	(0.0242)	(0.0442)	(0.0201)	(0.0666)
Ratio of Payout to Premium			0.123***	0.246***
			(0.0237)	(0.0405)
Ratio of Payout to Premium ^2			-0.0120***	-0.0243***
			(0.00263)	(0.00409)
Year 2006 Dummy	-0.0250**	-0.0269	-0.0395***	-0.0336
	(0.0111)	(0.0274)	(0.0107)	(0.0274)
Constant	0.0704***	0.165***	0.0771***	0.168***
	(0.00901)	(0.0172)	(0.00898)	(0.0173)
Marketing Restricted Sample	NO	YES	NO	YES
Observations	10,997	4,201	10,997	4,201
R-squared	0.014	0.035	0.034	0.058

Robust standard errors in parentheses — State Level Fixed Effects Included
*** p<0.01, ** p<0.05, * p<0.1 — Errors Clustered at the Village Level

Here $y_{i,t+1}$represents whether subject i purchases insurance at time $t+1$, and $P_{i,t}$ is a dummy variable that takes a value of 1 if person i receives an insurance payout at time t.[14] The sample is all buyers of insurance from 2005 and 2006, and I include a dummy (D_{2006}) that takes a value of 1 for purchasers in the year 2006 to control for time effects. Also, I only include purchasers who have weather insurance contracts available in their area in the following year.[15] These results are presented in Table 4, and Column 1 reports the baseline OLS results. It shows that receiving a payout is associated with a 9% increased chance of repurchasing insurance the following year, which means that those who receive an insurance payout are more than twice as likely to purchase insurance the following year as those who did not receive a payout. The dummy for 2006 is negative and significant, which is expected due to the minimum sum insured rules imposed in 2007.

One point of concern with these results is that there are many cases where there are multiple purchasers of insurance in a certain village in one year, and then zero in the next year. While this could be the result of people simply being unsatisfied with insurance, the large amount of villages that suddenly drop to zero purchasers is suspicious. As noted before, I don't know if BASIX marketed rainfall insurance in a particular village, or even if a certain village was visited by BASIX at all. For all the villages that had purchasers in one year and then none in the next year, it is quite likely that no BASIX representative visited the village, and therefore the customer didn't really have a chance to purchase the insurance. If this was the case it would make sense to exclude these villages from the

[14]It makes sense to assume that the error $\epsilon_{t,i}$ is correlated for the same person across time, as well as across people in a given year. Ideally, we would like to include individual fixed effects to account for individual heterogeneity. However, in order to exploit this variation we would need to look at customers who purchased insurance in both 2005 and 2006, and received payouts in only one of those years. Unfortunately, due to the very low repurchase rate, this results in very little variation and is therefore an unsuitable method of analysis.

[15]Basix's insurance coverage area varied somewhat from year to year. Results do not change significantly if all areas are included in the regression.

analysis, as the previous year's payout would have no way to possibly influence a customer's purchase decision.

In Column 2 I exclude villages that had no purchasers the following year from the analysis, creating what I call the 'Marketing Restricted Sample'. For instance, say village A had 10 purchasers in 2005, 13 purchasers in 2006, and 0 in 2007. In this case, the buyers from 2005 would be included in the sample since they obviously had opportunity to purchase the next year. However, the 2006 buyers would be excluded because I make the assumption that they didn't have the opportunity to buy in 2007. Restricting the sample this way results in a drop of the number of observations from 11002 to 4202, and causes the coefficient on receiving a payout to more than double to .22. This lends some credence to the argument that the omitted information about whether a village received marketing was downward biasing the results.

The coefficients generated in this restricted sample may be incorrect, as the decision to market to certain villages and not others is most likely not exogenous. If the marketing teams decided whether or not to market to certain villages based on the previous year's rainfall or experience with insurance then the results could be biased. For instance, assume that there were a number of villages that experienced a rainfall shock but received very low payouts, making them unhappy with insurance. If the marketing team knew this they may have decided to not market to as many of these villages, therefore censoring villages that received a payout but were likely to have few repeat buyers. Regressions that use previous years' payout characteristics to try to predict whether insurance is sold in a village the following year do not reveal any patterns that would suggest selection bias, but they may miss more subtle selection patterns. It is possible that the coefficient for the marketing restricted sample is upward biased and it therefore would be reasonable to regard the coefficients in Columns 1 and 2 as lower and upper bounds respectively.

The loss aversion theory suggests that peoples' reference points may change when they experience a perceived gain from an insurance payout. Whether or not a customer perceives an insurance payout as a gain may depend on the amount of the payout in relation to the premium paid. In Columns 3 and 4 I add two new continuous variables to the regression: the ratio of the payout received to the premium paid (which I will call the "payout ratio") and the payout ratio squared. The payout ratio has a positive and strongly significant effect, while the squared term is smaller and negative. This suggests that higher insurance payouts result in greater propensity to purchase the following year, but that the marginal effects flatten out for larger payouts. Also, the simple dummy of receiving a payout flips to negative, suggesting that small payouts have a negative affect on purchasing. In fact, payouts have a positive effect only when the payout ratio nears 1. As it makes sense that customers would need to receive a net profit on their insurance transaction to experience a reference-changing "gain", this result fits in well with the loss aversion model.

One may be concerned that the linear probability model may give biased estimates, especially since such a small percentage of the sample were repeat buyers. For a robustness check I also ran the regressions in Columns 3 and 4 using probit and logit specifications, and the results are very consistent with those obtained from OLS.[16] Also, clustering standard errors at the weather station level still yields highly significant coefficients.

As mentioned earlier, the dependent variable in this regression was generated by manually matching customers from one year to another, and therefore may be measured with error which could bias the coefficients downward. In order to get a feel for the potential magnitude of this error I run simulations where I assume that the BASIX data has been matched completely correctly, and then induce 'measurement error' by randomly changing the dependent variable of whether people purchased the following year or not. With the introduction of 10% matching errors (with an equal probability of a mismatch for buyer or non-buyers), the coefficient on receiving a payout in the full sample (Column 1) falls from .090 to an average of .072 over 1000 simulations. For the marketing restricted sample in Column 2, it drops from .222 to .178. In other words, if we assume 10% matching errors, then the estimated coefficients are likely to be underestimated by around 20%. It also may be possible that most of the error came from being unable to identify positive matches, possibly due to different members of a household signing the insurance contract from year to year. Repeating the above simulation but assuming that only people who were found not to have bought the next year could have been errors, the coefficients become underestimated by around 10%. While the exact form and structure of the matching errors cannot be known, it is likely that the reported coefficients are somewhat lower (in absolute value) than the true coefficients.

Overall, the results indicate that receiving an insurance payout correlates with a roughly 9-22% higher chance of repurchasing the next year compared with someone who purchased insurance but did not receive a payout. They also suggest that higher payouts lead to a greater chance of repurchasing, and that very low payouts may actually have a negative effect. While all these results are consistent with the theory of shifting reference points, they are also consistent with a number of other explanations, such as receiving payouts causing increased trust in the insurance company. The next section will attempt to empirically isolate some of these other mechanisms to see if they can explain the effects found in this section.

5 Alternative Explanations

While observing that people who receive insurance payouts are more likely to purchase insurance the following year is consistent with this paper's loss aversion theory, it is also consistent with many other explanations. In this section I attempt to isolate some of these other mechanisms to see whether they might instead be driving the results. I first consider the hypothesis that a rainfall shock as opposed to the insurance payout may cause people to be more likely to purchase the following year. To do this I look at villages in the first year they were offered insurance, and see if a rainfall shock the

[16] Results not shown.

previous year correlates with greater insurance take-up. On the contrary, I find that villages that had a rainfall shock the previous year were actually less likely to purchase insurance the following year, which provides strong evidence against the argument that weather as opposed to payouts are driving the main result.

I next consider the widely hypothesized suggestion that receiving insurance payouts would cause people to gain trust in the insurance company and learn about insurance, therefore making them more likely to purchase insurance in the future. To do this I assume that in order to gain trust in the insurance company or learn how insurance works, one would not have to receive a payout themselves; witnessing a neighbor receive a payout should also have the same effect. I therefore look for evidence of spillovers within a village and do not find evidence that witnessing a payout without actually receiving it yourself has a significant effect on the propensity to purchase the following year.

I also consider the possibility that payouts cause increased take-up due to direct wealth and/or liquidity effects as opposed to psychological effects. While I do not have data to empirically separate these possible mechanisms, I argue that due to the timing and circumstances of rainfall insurance payouts, wealth and liquidity are unlikely to play an important role.

Finally, I address the concern of unobserved marketing variation. While the effect of this omitted variable is admittedly difficult to measure, I argue it is unlikely to be driving the central results.

5.1 Direct Effects of Rainfall

Since most rainfall insurance payouts come at the same time as a rainfall shock, it is possible that the rainfall shocks themselves as opposed to the insurance payouts are what is driving increased take-up the following year. There is some evidence for this happening in developed markets, as Kunreuther et al. (1985) note that purchases of flood and earthquake insurance in the US spike after a recent event, even if people were not insurance customers before.

There are a number of theories that could explain this behavior. First, recent experiences with rainfall could change subjects' beliefs about the probability of a rainfall shock the following year. If there is actual autocorrelation of rainfall events or if the subject has limited knowledge about rainfall shocks, people may update their beliefs about shocks and therefore have more desire for insurance the following year. Alternatively, recently experiencing a rainfall shock could make shocks more salient, increasing the chance they will buy insurance the following year. Also, rainfall shocks may affect the wealth of the farmers. If farmers become poorer due to bad rainfall, CRRA utility would suggest that they would be even more risk averse the next year as a second shock would cause greater disutility.

I start by examining whether there is actual autocorrelation in the rainfall data. To test for autocorrelation, I create a panel of various rainfall indicators from 1961-2004 for each weather station. For each indicator, I run a regression of six lags of the variable on the current value, including weather station fixed effects. These results are presented in Column 1 of Table 5, with just the coefficient on

Table 5: First Order Autocorrelation of Weather Variables

	Fixed Effects	Arellano-Bond
	(1)	(2)
Total Rainfall	-0.106***	-.086***
	(.030)	(.021)
Phase 1 Rainfall	-.090***	-.075***
	(.030)	(.029)
Phase 2 Rainfall	-.018	-.026
	(.030)	(.028)
Phase 3 Rainfall	-.029	.007
	(.030)	(.028)
Would Have Been Payout	.023	.017
	(.030)	(.022)
Total Insurance Payout	-.0353	.004
	(.030)	(.028)
Weather Station Fixed Effects	YES	YES

Observation are years 1967-2004 for Fixed Effects Regression
Observation are years 1962-2004 for Arellano-Bond Regression
Fixed Effects regression contains six lags, Coefficient of First Lag Displayed
Arellano-Bond Regression contains one lag
Standard Errors are in Parentheses
*** p<0.01, ** p<0.05, * p<0.1

the first lag shown. While a fixed effects regression with a lagged dependent variable is not generally consistent, it will converge to the true value as $T \to \infty$. As T is relatively large (38), these estimates are likely to suffer from little bias. I also run a regression of the first lag using previous lags as instruments, using the methodology proposed by Arellano and Bond (1991), with results presented in Column 2. The results from both specifications are similar, and show a negative first-order autocorrelation in rainfall that appears to be driven by rains early in the season. The bottom two rows test for autocorrelation of rainfall shocks using the parameters of the 2005 insurance policy to determine shocks. "Would Have Been Payout" is a dummy variable that takes a value of 1 if the insurance policy of 2005 would have given a payout, while "Total Insurance Payout" is the size of this payout. By these measures, shocks do not appear to exhibit significant positive first-order autocorrelation.

This evidence casts doubt on the hypothesis that positive autocorrelation of weather events is driving increased insurance purchasing. It appears that total rainfall is actually negatively autocorrelated, while shocks (which are proxied by the insurance contract giving a payout) do not appear to be correlated at all.

Even if there is no positive autocorrelation of rainfall, there may be other aspects about experiencing a shock that result in people having a higher propensity to purchase insurance. In order to look at the results of weather separately from the effects of insurance, I analyze how previous weather events affected insurance purchase decisions in the first year that insurance was offered to BASIX customers, which was 2005. To accomplish this, I first aggregate the purchasing data to the village level and then test to see whether villages that experienced a rainfall shock in 2004 had more insurance purchasers in 2005 than villages who did not experience a rainfall shock. A shock is defined using each location's

Table 6: Effect of Shocks on Purchasing

	Dependent variable is number of buyers in 2005			
	(1)	(2)	(3)	(4)
Would Have Been Payout in 2004	-3.843***	-4.592***	-5.045**	-3.788*
	(0.987)	(1.039)	(2.173)	(1.898)
Ratio of 2004 Payout to 2005 Premium			4.365	-0.755
			(4.610)	(5.543)
Payout Ratio Squared			-1.991	-0.279
			(1.814)	(2.064)
Constant	8.001***	0.651	7.985***	1.015
	(0.714)	(6.341)	(0.713)	(6.494)
Weather Station Constants	NO	YES	NO	YES
Observations	733	733	733	733
R-squared	0.073	0.094	0.075	0.097

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

Observations weighted by quality of rainfall data
Errors Clustered at Weather Station Level
All Regressions Include State Fixed Effects

insurance policies in 2005: If insurance would have paid out in 2004 based on the structure of the 2005 weather policy, this is deemed a rainfall shock. As the quality of the rainfall data is related to the amount of nearby weatherstations, I weight the observations based on the number of nearby rainfall stations.[17] Also, I create a hypothetical payout ratio, similar to the "Ratio of Payout to Premium" variable presented in Table 4. This is the ratio of the amount that the 2005 policy would have paid out in 2004 divided by the premium of the policy.

The results of this regression are presented in Table 6.[18] Column 1 presents the baseline regression, which shows that villages that experienced a rainfall shock in 2004 actually had an average of 3.8 *fewer* purchasers in 2005. One worry with this regression may be that since the insurance policies and rainfall patterns of each location are different, the definition of a shock may vary from one place to another. Therefore, the estimates may be improved with the inclusion of location and policy-specific covariates, which I title "Weather Station Constants". In Column 2 I add controls for the historical average rainfall, historical rainfall standard deviation, the policy premium in 2005, historical average payout of the policy, and the percentage of historical years there would have been a payout. Note that all the "historical" data is calculated from 1961-2000. With the addition of these controls, the

[17] The APHRODITE weather data provides information about how many local weatherstations contributed to a certain rainfall reading. Since some of the rainfall observations are likely to be more accurate than others, I weight them according to accuracy. If there are no rainfall stations contributing to the APHRODITE data within a $.75^o$x$.75^o$ grid around the desired BASIX weather station, the observation is given a weight of 1. If there is a least one weather station in this $.75^o$x$.75^o$ grid, the observation is given a weight of 1.5. If there is a rainfall station within the $.25^o$x$.25^o$ grid, the observation is given a weight of 2. The weighted results to not differ significantly from the unweighted results.

[18] Note that while it is reasonable to think that village-specific characteristics (such as village size) may have an effect on village-level insurance take-up, village-level co-variates are not included in the regression. When the regressions are run with the village characteristics from the 2005 Indian census, the coefficients of interest do not change significantly. Also, most village-level characteristics had insignificant coefficients, with the exception that a more literate population was correlated with higher takeup. Since village-level coefficients were only available for around 50% of the villages, these variables are not included in the main specifications.

coefficients on having a rainfall shock in 2004 remains negative, and even decreases slightly.

Following previous results that suggest that the size of the insurance payout is important, in Columns 3 and 4 I include variables for the severity of the shock in 2004 using the ratio of the hypothetical payout to the premium (the payout ratio) and the payout ratio squared. In both specifications these variables are insignificant, suggesting that most of the variation in purchasing in 2005 is explained by our binary shock variable.

The main conclusion to be drawn from these regressions is that the data does not support the hypothesis that bad weather induces people to purchase insurance in the following season. If anything, it seems to decrease insurance purchases. I can only speculate on the reasons for this; it may be due to the fact that people recognize the actual negative autocorrelation of rainfall, or it may be that the rainfall shocks decrease the available liquidity to purchase insurance the following year. Regardless, this data provides relatively convincing evidence that the direct effect of weather is not causing people who receive insurance payments to purchase again the following year.

5.2 Trust, Learning, and Spillover Effects

It is also possible that the propensity to purchase insurance after receiving a payout results from learning about insurance and trusting the insurance company, as opposed to being a direct result of the payout. In order to separate the effects of trust and learning from that of receiving the payout, I make the assumption that if trust and learning are playing an important role in causing people to purchase insurance after they have received a payout, then we should be able to see a positive spillover effect of payouts within the village.[19] This is because one shouldn't need to actually receive a payout to gain the effects of trust and learning, as someone who witnesses a payout gains all the same information as someone who receives a payout. But witnessing a payout would not give the psychological effect of gain from the insurance company utilized in the loss aversion model.

To perform this analysis I aggregate all buyers to the village level, but divide them into two types: repeat buyers and new buyers, where repeat buyers are people who purchased insurance the year before. I then regress the number of each type of buyer on payout statistics and the total number of buyers in the previous year. When there is an insurance payout in the previous year, most of the repeat buyers the following year received money from the insurance company, while new buyers didn't receive anything.[20] These results are presented in Table 7.

In order to compare results with the main specification in Table 4, I again provide a dummy for whether there was a payout in the village along with a quadratic effect of the ratio of payouts to the premium. When aggregating the village data I used the mean of the payout ratios in the village to

[19] If people can only gain trust and learning by actually receiving a payout themselves, then then data gives us no way to separate trust and learning from other possible mechanisms of receiving a payout.

[20] Some buyers may not have received money if they bought one phase of the insurance policy but one of the other phases paid out. This happened in 427 cases, and removing these individuals does not change there results.

Table 7: New Buyers In A Village

	Dependent Variable is the Number of Buyers in a Village the Following Year					
	Panel A: All Villages			Panel B: Villages With At Least 1 Repeat Buyer		
	Total Buyers	New Buyers	Repeat Buyers	Total Buyers	New Buyers	Repeat Buyers
	(1)	(2)	(3)	(4)	(5)	(6)
Was Payout in Village	-0.334	0.376	-0.709**	3.353	5.799*	-2.446***
	(2.252)	(2.121)	(0.338)	(3.275)	(3.413)	(0.433)
Mean Ratio of Payout to Premium	1.319	0.481	0.838*	0.522	-2.131*	2.653***
	(1.330)	(1.050)	(0.463)	(1.094)	(1.095)	(0.168)
Mean Payout Ratio Squared	-0.135	-0.0617	-0.0729	-0.0832	0.166	-0.249***
	(0.141)	(0.106)	(0.0539)	(0.120)	(0.114)	(0.0167)
Number of Buyers in Village	0.131***	0.0795*	0.0512***	0.197***	0.0959	0.101**
	(0.0480)	(0.0434)	(0.0178)	(0.0530)	(0.0596)	(0.0427)
Year 2006 Dummy	-2.994*	-2.738*	-0.256	-5.231	-5.241*	0.0105
	(1.661)	(1.448)	(0.236)	(3.284)	(2.929)	(0.514)
Constant	3.445***	3.301***	0.144	10.48***	10.00***	0.483*
	(1.140)	(1.065)	(0.136)	(1.599)	(1.614)	(0.249)
Observations	1534	1534	1534	459	459	459
R-squared	0.061	0.047	0.118	0.084	0.069	0.285

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1
Errors clustered at the weather station level. All regressions include state fixed effects
Data is aggregated to the Village Level
Includes all villages in 2005 and 2006 where there was insurance coverage the following year

create a payout ratio for the village.[21] The overall results of the table tell a clear story: payouts drive repeat buyers but not new purchasers, showing few spillover effects. Column 3 shows how payouts affect the number of repeat buyers the next year, and the results are very consistent with the baseline results from Table 4. A dummy for whether there was any payout is negative and significant, but the payout size has a positive effect. This suggests that low payouts have a marginally negative effect on the number of repeat purchasers, but this effect flips to positive as the size of the payout ratio increases above 1. Column 2 shows the effect of payouts on new buyers in a village. Here all the payout coefficients are insignificant, but due to large standard errors I cannot reject that they are the same as the effects on repeat buyers.

In Panel B I restrict the analysis to villages that had at least one buyer the year after insurance outcomes, creating a sample analogous to the 'Marketing Restricted Sample' in Table 4. The logic behind this is that if a village had zero buyers it is likely that insurance was not marketed in the village in that year, and therefore customers did not have an opportunity to purchase insurance. Restricting the data set in this way gives a much clearer pattern. Column 6 now shows much stronger effects of payouts on repeat buying, though the pattern is the same as in Column 3. Small payouts have a negative effect, while increasing the payout ratio increases repeat buying. The squared term on the payout ratio is now negative and significant, indicating that high payout ratios have diminishing effects.

These coefficients are now all significantly different from the coefficients for new buyers found in

[21] The results are not sensitive to using the mean, and are very similar using the median, maximum, and mode.

Column 5. In fact, the coefficients in Column 5 flip signs, suggesting that payouts have the opposite effect on people who did not receive payouts. These results suggest that low payouts actually induce more new buyers, but that these effects decrease and then turn negative as the payout in the village increases. This effect isn't consistent with any of the theories I have advanced, but is especially inconsistent with the hypothesis that people are purchasing insurance after receiving a payout due to the effects of trust and learning, as we have seen that higher payouts increase peoples' propensity to purchase again the next year.

One important clarification of these results is that most of the potential "new buyers" living in a village that had experienced payotus would have also experienced uninsured rainfall shocks during the same season. Therefore it may be possible that there are effects of trust and learning, but they are outweighed by opposite effects of the weather. As we saw in the previous section, rainfall shocks tend to have a negative effect on insurance demand, so the (lack of) evidence of spillovers may be a result of a more complex interaction between trust/learning and direct effects of weather.

Overall, these results do not support the hypothesis that trust, learning, or any other effects of simply witnessing insurance payouts are driving increased purchasing. While it is possible that our measurements of spillovers are too crude and miss more subtle effects, it is telling that there is no sign of spillovers in villages that received the largest payouts. Since we do not see these spillover effects, this provides further evidence that increased purchasing of insurance is instead driven by the actual reception of money from the insurance company, and is consistent with the proposed loss aversion model.

5.3 Direct Effects of Payouts on Wealth and Liquidity

The previous two sections discount the possibility that trust, learning, or weather effects are driving the result that receiving an insurance payout is correlated with purchasing insurance the following year. This points to the actual reception of money from the insurance company as being the driving force behind greater takeup. However, this paper's model of loss aversion is not the only explanation that could explain why an influx of money could drive greater insurance uptake. Instead, one might think that receiving an insurance payout could directly affect choices the next year due to its effects on wealth and liquidity. For instance, if insurance were a normal good then increased wealth would result in greater insurance demand.[22]

While the BASIX data set does not offer the opportunity to test the direct effects of a cash payment separately from an insurance payout, there are a number of reasons why it is unlikely that wealth or liquidity effects are driving the results. Most importantly, insurance payouts are given in the context of a rainfall shock, which would most likely result in a loss of income. It may help to recall that the empirical results are being driven by variation in rainfall across locations, not by levels of insurance

[22] This is consistent with the empirical findings of Cole et al. (2010).

within a village. Therefore, for wealth effects to be driving the results, one would need to think that experiencing an insurance payout in the context of a rainfall shock resulted in people becoming wealthier than those people who didn't experience a shock at all. Given the fact that most buyers bought a relatively low amount of insurance coverage relative to their incomes, experiencing a rainfall shock, even when insured, would likely decrease future wealth. Therefore, wealth effects seem like a poor explanation as to why receiving payouts spur future insurance sales.

If people who received insurance payouts had a decrease in wealth it is also unlikely that receiving the insurance payout would increase their liquidity the next season. Insurance payments were generally made in January, while people had the opportunity to purchase insurance for the next season only in May. It is doubtful that these payments would have a lasting enough liquidity effect to influence insurance buying decisions five months later.

While I can't provide direct empirical evidence against the hypothesis that insurance payments drive increased take-up due to wealth or liquidity effects, given the structure and timing of insurance payments this explanation seems extremely unlikely.

5.4 Omitted Marketing Intensity

As mentioned earlier, the data set does not contain the exact marketing practices that BASIX undertook in each village in each year. If the intensity of marketing was correlated with both previous years' insurance payouts and current years' sales, this omitted variable could be biasing the results. For instance, assume that the marketing staff at BASIX think that people who have just received a payout are more likely to repurchase insurance. In this case, as the marketing team has limited resources, it may make sense for them to direct these resources towards the area of highest return, which would be people who have already received payouts. If this was the case, the increased take-up rates from people who received payouts could simply result from increased marketing attention from the BASIX team.

While the results could be picking up some of this effect, there are a couple of reasons I believe it is unlikely to be a significant factor. First, regressions of observable marketing factors (such as a dummy of whether there were any purchasers in the village) do not show any significant correlations with payouts. Next, the BASIX marketing staff claim to not give any special marketing treatment to previous payout recipients.[23] As they are trying to build long-term business, BASIX claims that they do not change their marketing practices for villages that have recently received a payout. Finally, if BASIX targeted payout recipients and they didn't really have a higher tendency to purchase, one would think that the marketing team would quickly learn that this strategy wasn't effective and would stop it. While I only observe two marketing cycles and erroneous beliefs could survive throughout this short time span, it is telling that the effect of payouts on take-up is greater in 2006 than 2005,

[23] Conversation with Sridhar Reddy, Assistant Manager for Insurance at Basix, Jan 09.

suggesting that the effect is increasing over time.[24] If it was caused by erroneous expectations of the marketing team, we would expect the effect to decrease over time. Overall, while I must accept the possibility that increased marketing is driving the results, I regard it as unlikely.

6 Conclusion

After receiving an insurance payout, customers of rainfall insurance in India are 9-22% more likely to purchase insurance again the next year. This behavior seems to be driven by actually receiving the money from the insurance company, and is consistent with a loss aversion model where previous insurance gains shift the subject's reference point. In this model, a subject views future insurance premiums as deductions from his previous gains, as opposed to a true loss. Therefore, after receiving an insurance payout, future insurance purchases are more attractive. While there are other possible explanations for this phenomenon, the BASIX customer data does not provide support that any of these other possible mechanisms are driving the results.

First, direct effects of a rainfall shock do not appear to drive insurance purchases. Looking at villages in their first year of insurance availability, I find that locations that experienced a rainfall shock the year before are actually less likely to purchase insurance. Next, I do not find evidence that repeat purchasing is driven by trust, learning, or any other effect that one would expect to spill over to other members of the community. Taken together, these results point toward the actual reception of money from the insurance company as the dominant driver of repeat purchases. While this does not constitute direct evidence of the loss aversion model presented in the paper, it does provide empirical evidence that is consistent with the predictions and mechanisms of the model.

This study brings to light a number of questions that would be ripe for future research. First of all, it would be interesting to understand whether insurance payouts have long-term effects on future purchases, and also whether payouts continue to have similar effects for people who have years of experience with insurance. To answer these questions one would need a data set with a longer time frame. Also, a longer data set could shed further light onto the question of whether customers learn about insurance over time. It is possible that people need a few years of experience with insurance to really learn about the product and gain trust in it, which would explain why this paper fails to see any spillover effects.

These results point to a number of policy recommendation for the Indian rainfall insurance market, and possibly for insurance markets in general. One of the main arguments made for the slow adoption of insurance in India is that people do not understand insurance and do not trust the insurance companies. If trust and learning were the crucial determinant of insurance adoption, then incentives could be given to encourage early adoption and over time as people witnessed and experienced payouts we would expect insurance adoption to grow. This paper fails to find any evidence of increased trust

[24] Results not shown.

and learning driving insurance decisions, which suggests that incentivizing early adopting is unlikely to quickly spur insurance take-up.

Historical evidence (as in Kunreuther et al. 1985) has suggested that an effective policy to spur insurance markets would be to target areas that have recently experienced a large shock. This paper does not support this notion in the case of rainfall index insurance in India, as places that recently experienced a shock were less likely to purchase insurance.

Instead I suggest that the mechanism that drives increased purchases after a payout is the feeling of winning money from the insurance company. This result is not entirely helpful in terms of policy prescriptions. If the propensity to purchase insurance after receiving a payout is truly due to the reception of money, in order to maintain customers the insurance company would have to give out significant payouts each year. Since this would result in losses for the insurance company, such a scheme would not be sustainable in the long run.

With relation to the future of rainfall index insurance in India, one stark result is that the raw numbers of continuing customers of insurance are very low, calling into question the sustainability of the product. Even among people who received payouts in excess of twice their premium in 2006, only 18% bought again in 2007. With the proportion of repeat buyers so low, one would have to assume that many people are not satisfied with their experience of insurance, which suggests that the product or marketplace will need to evolve in order to survive.

One factor to note is that this study looks at the first major scale-up of rainfall insurance in the world. Rainfall insurance is still a young product, and is still evolving to meet the needs of customers. One particular point of attention is the massive loading on most policies offered. As we saw in Table 2, the many BASIX insurance policies had premiums of up to six times the actuarially fair rate. With premiums this high, it is unsurprising that people are not signing up. Also, one may argue that the correlation between insurance payouts and crop outcomes were less than ideal in these early products. Around the world, index insurance policies are constantly evolving to better correlate with crop outcomes and avoid basis risk. While this study predicts that rainfall insurance in the form of BASIX's policies from 2005-2007 are likely to fail, it is quite possible that innovations in products and pricing can create an insurance product that better meets the needs of small scale farmers.

References

Arellano, Manuel and Stephen Bond, "Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations," *The Review of Economic Studies*, 1991, *58* (2), 277–297.

Bryan, Gharad, "Ambiguity and Insurance," Nov 2010. Technical Report, mimeo.

Cole, Shawn A., Xavier Giné, Jeremy Tobacman, Petia Topalova, Robert Townsend, and James Vickery, "Barriers to Household Risk Management: Evidence from India," *Harvard Business School Working Paper*, 2010, *09* (116).

Giné, Xavier, Robert Townsend, and James Vickery, "Patterns of Rainfall Insurance Participation in Rural India," *World Bank Economic Review*, October 2008, *22* (3), 539–566.

Gomes, Francisco J., "Portfolio Choice and Trading Volume with Loss-Averse Investors," *Journal of Business*, March 2005, *78* (2), 675–706.

Hess, Ulrich, "Innovative financial services for India, monsoon indexed lending and insurance for smallholders," *ARD Working Paper*, 2004, *09.*

Hill, Ruth and Miguel Robles, "Flexible Insurance for Heterogeneous Farmers: Results from a Small Scale Pilot in Ethiopia," 2010. Technical Report, mimeo.

Kahneman, Daniel and Amos Tversky, "Prospect Theory: An Analysis of Decision under Risk," *Econometrica*, 1979, *47* (2), pp. 263–292.

Köszegi, Botond and Matthew Rabin, "A Model of Reference-Dependent Preferences," *Quarterly Journal of Economics*, 2006, *121* (4), 1133–1165.

_ **and** _ , "Reference-Dependent Risk Attitudes," *The American Economic Review*, 2007, *97*, 1047–1073(27).

Kunreuther, Howard, Warren Sanderson, and Rudolf Vetschera, "A behavioral model of the adoption of protective activities," *Journal of Economic Behavior and Organization*, 1985, *6* (1), 1 – 15.

Platteau, Jean-Phillipe, "Mutual Insurance as an Elusive Concept in Traditional Rural Communities," *Journal of Development Studies*, 1997, *33* (6), 764–796.

Pratt, John W., "Risk Aversion in the Small and in the Large," *Econometrica*, 1964, *32* (1/2), pp. 122–136.

Thaler, Richard H. and Eric J. Johnson, "Gambling with the House Money and Trying to Break Even: The Effects of Prior Outcomes on Risky Choice," *Management Science*, 1990, *36* (6), pp. 643–660.

Townsend, Robert M., "Risk and Insurance in Village India," *Econometrica*, 1994, *62* (3), pp. 539–591.

7 Appendix

7.1 Solving the Model for the General Case

Section 2 of the paper solved our reference-dependent model for a specific case of $r_1 = 0$, $f(r_1, 1, 0) = r_1$. This is the simple situation where the reference point starts at zero and we assume that if someone has bought insurance but does not receive a shock then their reference point will not change. In this Appendix I solve for the case of a general r_1 and $f(r_1, 1, 0)$, though still keeping the restriction that $-X < f(r, I, S) < 0$, as this condition ensures nonlinear (and therefore interesting) utility functions over the relevant range. Remember, $B_2(f(r_1, 1, 1)) > B_2(f(r_1, 1, 0))$ means that the benefit of buying insurance in the second period is greater for an insurance customer in period one who has received an insurance payout than for a customer who did not receive a payout.

Proposition 2. $B_2(f(r_1, 1, 1)) > B_2(f(r_1, 1, 0))$ iff

$$\begin{cases} f(r_1,1,0) > f(r_1,1,1) > \frac{p-1}{p} f(r_1,1,0) - (1+\lambda)X & if \quad -(1+\lambda)pX < f(r_1,1,0) < 0 \\ f(r_1,1,0) < f(r_1,1,1) < -\frac{p}{(1-p)}(f(r_1,1,0) + (1+\lambda)X) & if \quad -X < f(r_1,1,0) < -(1+\lambda)pX \end{cases}$$

Proof. To prove Proposition 2 , we must proceed case by by case. Start with the case where $-(1+\lambda)pX < f(r_1, 1, 0) < 0$, which means that customers who don't get an insurance payout in period 1 have a reference point in period 2 greater than the insurance premium. In this case the second period benefit for a customer who does not receive an insurance payout is:

$$B_2(f(r_1,1,0)) = -\lambda\beta pX + (1-p)(\alpha-\beta)f(r_1,1,0) \tag{9}$$

For those who did receive a payout, benefits are:

$$\begin{cases} B_2(f(r_1,1,1)) = -\lambda\beta pX + (1-p)(\alpha-\beta)f(r_1,1,1) & if \quad -(1+\lambda)pX < B_2(f(r_1,1,1)) < 0 \\ B_2(f(r_1,1,1)) = -(1+\lambda)\alpha pX + p\beta X - p(\alpha-\beta)f(r_1,1,1) & if \quad -X < f(r_1,1,1) < -(1+\lambda)pX \end{cases} \tag{10}$$

Combining Equations 9 and 10 proves the first case of Proposition 2. Turning to the case where $-X < f(r_1, 1, 0) < -p\lambda X$, the benefits for people who did not receive a payout are

$$B_2(f(r_1,1,0)) = -(1+\lambda)\alpha pX + p\beta X - p(\alpha-\beta)f(r_1,1,0) \tag{11}$$

The benefits for those who did receive a payout are:

$$\begin{cases} B_2(f(r_1,1,1)) = -\lambda\beta pX + (1-p)(\alpha-\beta)f(r_1,1,1) & if \quad -(1+\lambda)pX > f(r_1,1,1) > 0 \\ B_2(f(r_1,1,1)) = -(1+\lambda)\alpha pX + p\beta X - p(\alpha-\beta)f(r_1,1,1) & if \quad -X < f(r_1,1,1) < -(1+\lambda)pX \end{cases} \tag{12}$$

Combining Equations 11 and 12 proves the second case of Proposition 2.

$\underset{r}{Max}\, B_2(r) \rightarrow r = -(1+\lambda)pX$

In words, this means that a customer achieves the greatest benefit from insurance when the reference point is equal to the negative of the premium of the insurance policy.

For an arbitrary reference point r, the benefits of insurance can be calculated by combining Equations 2 and 3.

$$\begin{cases} B(r) = -\lambda p\beta X - r(\beta + \beta\alpha + p(\beta-\alpha)) & if \quad -(1+\lambda)pX < r \\ B(r) = -(1+\lambda)\alpha pX + p\beta X + rp(\beta-\alpha) & if \quad -(1+\lambda)pX > r \end{cases}$$

In the first case, $B(r)$ is maximized at the boundary of $r = -(1+\lambda)pX$ since $(\beta+\beta\alpha+p(\beta-\alpha)) > 0$. In the second case $B(r)$ is maximized at the boundary of $r = -(1+\lambda)pX$ since $(\beta-\alpha) > 0$. $\square$