The effect of individual agricultural insurance on transfers by smallholder farmers in informal risk-sharing networks: empirical evidence from Tigray, Ethiopia

Qhelile Nyathi

Linacre College

University of Oxford

A thesis submitted for the degree of

MSc by Research in Statistics

Trinity 2014

Acknowledgements

My deepest gratitude goes to Dr Karlijn Morsink and Dr David Steinsaltz. Their contribution to my learning has been invaluable. I am also grateful for my fiance, Xolani Ndlovu and my family who continue to support me through distance and time.

Abstract

We investigate the effect of formal individual agricultural insurance on transfers by farmers in informal risk sharing networks using an experiment conducted in rural Tigray in Ethiopia. Both index and indemnity insurance are investigated for a holistic picture. However, we focus on index insurance, particularly when there is basis risk.We employ a multi-stage sampling design in a framed field experiment, simulating three possibilities, namely, insurance availability, insurance take-up and insurance rejection.

Based on empirical considerations, we treat transfers as categorical and use multilevel analysis to fit a Cumulative Link Mixed-effects Model (CLMM). Thus we compare the odds of transfers under the three possibilities; insurance availability, insurance take-up and insurance rejection to transfers in a control in which formal insurance is unavailable. Having explored how transfers change in response to the three possibilities, we tentatively explore the underlying motives.

We find that farmers do not offer compensation for basis risk. When there is basis risk, farmers are 48% less likely to make transfers. This non-compensation for basis risk is worst for farmers in prescribed groups. Assuming socio-economic-dependent factors to be the underlying motives, we posit that the resultant economic incentives motivate farmers to punish their partners for their insurance decisions.

Providers of micro insurance in low-income farming markets will find the empirical findings useful in that they show the need for careful evaluation when designing agricultural insurance interventions aimed at realizing a mutually beneficial relationship between index insurance and informal risk sharing. The findings also highlight the need for caution in prescribing farmers into groups as a condition for offering insurance.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Background and objective

Farming is a risky business, more so for smallholder farmers in developing countries who practice rain-fed agriculture, have limited access to financial risk markets and are without governmental or public insurance. The situation is not completely dire though, because farmers support each other. Friends, family and neighbours in these farming communities form networks, which they use as mechanisms of pooling and managing risk. Development literature terms these *informal risk sharing networks* [??????]. Members transfer cash or in-kind assets, which we call *transfers* [??], in response to each other's needs. A member is said to be in need when he or she suffers a loss or experiences a shock. This could be in the form of a bad harvest, the death of livestock or even the loss of a loved one.

Although informal risk sharing provides informal insurance, this cover is often incomplete due to the fact that these networks consist of farmers in the same geographical unit, and weather related agricultural shocks affect people living in close proximity at the same time. These are called *aggregate shocks*, and their covariate nature limits the capacity of farmers in the same network to provide complete insurance to one another. In addition, farmers are also exposed to individual risk, which is referred to as *idiosyncratic shock*. These types of shocks are often uncorrelated within a group of farmers. Therefore, informal risk sharing, which can be thought of as informal insurance, is more effective at offering protection against idiosyncratic losses [**??**].

Formal insurance pools risk from different geographical units and over a much wider pool, hence is effective at covering aggregate shocks. This works particularly well in developed insurance markets, but less so in underdeveloped financial markets due to information asymmetries. In low-income markets, development practitioners and insurers traditionally provide indemnity insurance. This is a type of insurance that protects against a specific peril whereby the insured files a claim and the insurer assesses the loss prior to making a payout. In a developing country context, due to information asymmetries, indemnity insurance is riddled with problems of adverse selection, moral hazard and fraud. Additionally, the verification and payout procedure makes it too slow for smallholders' needs. Encouragingly, the last decade has seen the introduction of weather-based index insurance to solve the aforementioned problems [**??**].

The distinguishing feature of index insurance is that a payout is triggered by an external index, such as rainfall, humidity, temperature, or average yield, which can be based on directly measured historical yields, or satellite-based predictions of average yields. The external index is correlated with the individual insured farmer's yield, yet independent of his/her individual actions. Hence, this provides a solution to the major challenges of agricultural indemnity insurance. Firstly, it prevents moral hazard because a farmer cannot influence an index that is based on the weather. Secondly, it prevents adverse selection because the source of the risk is exogenous to the farmer, for example, the level of rainfall. Lastly, because an index triggers a payout, there is no need for the insured to file a claim and this minimises the time and cost associated with loss adjustment so it delivers a more timeous payout than indemnity insurance.

Yet, index insurance is not a panacea. It introduces a different problem, such as when the weather is poor enough for farmers to incur a poor crop and financial loss but not poor enough to trigger a payout. Conversely, it is possible for a policyholder to receive an index insurance payout but not experience a loss. The resulting imperfect correlation between the agricultural production loss and the index insurance payout is called *basis risk* [**??**]. Notwithstanding that there may be other risk factors causing the loss, for example pest infection, the existence of basis risk due to the recorded weather failing to trigger a pay-out makes index insurance risky and decreases its take-up by low-income farmers [**?**]. Insurers continue to improve product design to address this, and infrastructure and technology are continuously being improved to obtain more accurate weather information [**?**].

In the context of agriculture, aggregate and idiosyncratic shocks will certainly happen, but who suffers a shock and when a shock will be experienced is uncertain. Therefore, the underlying motives behind transfers in informal risk sharing networks are based on both preference-dependent factors; like risk preferences, inequality aversion and altruism, and on socio-economic incentives; namely sanctions and reciprocity [**??**]. However, in this thesis, we focus only on underlying transfer motives driven by socio-economic incentives. This is because introducing formal insurance alters only the socio-economic incentives and not the farmers' intrinsic preferences and so sanctions and reciprocity become more salient in determining transfer behaviour.

Therefore the choice to focus only on socio economic dependent factors when linking the transfers to their likely underlying motives is both a strength and a weakness of this study. The latter because preference-dependent factors may offer alternative underlying motives. The former in that we can more clearly disentangle the incentive-driven factors. In order to disentangle the pure effect of one farmer's insurance purchase decision from the motive of reciprocity we play a one-shot game, which means that for each pair we restrict the transfers to one season and one direction, hence a transfer can only flow from one farmer, who we will call the *donor* to another farmer, who we call a *recipient.* The one-shot game means that a recipient cannot reciprocate a donor's transfer. Therefore, a donor's transfer can only be motivated by sanctions and not reciprocity. Sanctions can be negative or positive; the former being punishment and the latter being a reward. We further prevent reciprocity by pairing farmers anonymously in the risk-sharing game of our experiment.

The primary aim of this study is to understand how farmers' insurance decisions change the transfer behaviour of other farmers in their risk-sharing network. That is, whether transfers increase, decrease or remain unchanged when insurance is taken-up or rejected in comparison to transfers with informal risk sharing when insurance is not available. Accordingly, we ground the observed changes in transfers to the possible underlying motives of punishment and reward. Secondary to investigating whether farmers in pre-existing networks who normally share risk informally, punish or reward those who take up or reject formal insurance, we investigate whether this is so for farmers from different informal risk sharing networks.

Formal insurance can therefore only be benign or welfare improving to a network depending on its impact on transfer behaviour. If index insurance is taken up and there is no aggregate shock to trigger a pay-out then a recipient farmer has an increased need for a transfer relative to a recipient farmer in the control treatment without any insurance. If index insurance impacts transfer behaviour positively, the amount of transfers will either increase in response to an increased need for a transfer or remain unchanged when index insurance is introduced, in which case index insurance and informal risk sharing are *supplements*. Therefore this study defines supplementarity as a *warranted* increase in the amount of transfers and replacement as a *warranted* decrease in the amount of transfers. We use warranted to highlight the need-based nature of transfers. The introduction of agricultural insurance in developing countries is assumed to improve welfare, when informal risk-sharing provides incomplete insurance against shocks and the formal insurance provides protection against aggregate shocks [**???**]. Meanwhile, recent theoretical results show that index insurance and informal risk-sharing are complements when index insurance carries basis risk and provides protection against the aggregate components of a shock while informal risk-sharing covers the idiosyncratic components [**??**]. However, there exists little empirical evidence on this complementarity. Through examining supplementarity, a more narrowly defined mutually beneficial relationship between index insurance and informal risk sharing, this study aims to make a contribution to this growing literature.

Existing studies on the interaction between index insurance and informal risk sharing assume take-up of insurance and enforceability of the informal risk-sharing network. Assuming enforceability necessarily presupposes transparency (or a degree of information asymmetry within the network). It assumes farmers in the network know each other's decisions. Take-up of insurance changes a farmer's risk-profile, thus it changes the recipient's level of need and the impetus for a donor farmer to make a transfer. In reality farmers may or may not take up index insurance when it is introduced. Furthermore farmers in the network may or may not know whether their counterparts have taken up insurance or not. Without full information on each other's decisions, there is ambiguity in the recipient's level of need for a transfer, which distorts the need-based motivation for transfers and may cause farmers to respond more to their partner's decisions (rather than the effect of their decisions).

Index insurance that carries basis risk increases the uncertainty associated with a farmer's income. When a farmer experiences basis risk the network is expected to respond by increasing or maintaining transfers. This rationale of complementarity is widely (mis)used by development practitioners and insurers who are increasingly offering index insurance on condition of farmers being members of a "group." In such cases farmers are prescribed into networks. Studies have looked at complementarity in pre-existing groups [**??**], but the extent to which informal risk-sharing and index insurance are complementary in prescribed groups is unknown. Although this study looks at supplementarity rather than complementarity per se, this study aims to contribute some nuance so that necessary caution is taken in agricultural insurance interventions.

Therefore the study investigates the effect of formal individual agricultural insurance on transfers by farmers in informal risk sharing networks. Both index and indemnity insurance are investigated for a more holistic picture but the focus is on index insurance.

The study examines the interaction between index insurance and informal risk sharing in a framed field experiment that more closely resembles the reality that insurance is available in the market but take-up is not guaranteed. The study further examines the change in transfers when the insurance introduced is taken up and when it is rejected. Finally the study compares these changes in networks that pre-date the introduction of index insurance (pre-existing) to networks that farmers are prescribed into, in order to access insurance.

Having explored how transfers change in response to a recipient farmer's insurance decision, we then tentatively explore the motives behind the change. Based on theory of motives for transfers in risk-sharing networks, we posit that payoff distributions create economic-incentives that motivate donors to either reward or punish their partners for their insurance decisions.

Therefore the study addresses the following questions:

(a) How do transfers in prescribed groups and pre-existing small holder farmers' risk-sharing networks change when either indemnity or index insurance is *introduced*?

(b) How do transfers in prescribed groups and pre-existing small holder farmers' risk-sharing networks change when either indemnity or index insurance is introduced *and taken up*?

(c) How do transfers in prescribed groups and pre-existing small holder farmers' risk-sharing networks change when either indemnity or index insurance is introduced *but rejected*?

This research is important from a development policy perspective, it helps us understand how, through manipulating economic-incentives, agricultural insurance interventions impact transfer behaviour in informal risk-sharing. The research is also exciting due to the researcher's interdisciplinary background in African studies, which adds critical reflexivity and relativity often lacking from a purely economic analytical lens. This research makes a novel contribution to scholarship on the nexus between formality and informality in smallholder farmers' risk management strategies through its empirical findings. Furthermore the research pays attention to the experimental and sampling design so as to carefully apply mixed-effects or multilevel statistical analysis methods to the data. These methodological components are often overlooked in social science experimental studies.

In the rest of this chapter we review related literature. In Section 1.3 we use the literature to inform predictions that can be expected from the research design. In Chapter 2 we elaborate on the research design, including the sampling design and the framed field experiment [?]. In Chapter 3 we present the data, exploring its quality, structure and distributions so as to inform the selection of a suitable statistical analysis method and modeling approach. Our data cannot be assumed to be normal. Furthermore, its hierarchical nature results in correlation between errors, hence in Chapter 4 we discuss the Generalized Linear Model and the Generalized Linear Mixed Model. However, due to the multimodality of the data, instead of a GLMM approach we settle on analysing the data as categorical. This is the subject of Chapter 5, where we fit Cumulative Link Mixed-effects models (CLMM) and present the results. Chapter 6 concludes.

1.2 Literature review

Agriculture provides 70% of Africa's full time employment, one third of total GDP, and 40% of total export earnings [?]. However, smallholder farmers remain vulnerable to weather risks and other shocks. Agricultural insurance, used as part of a broad agricultural risk management strategy can increase smallholder famers' financial resilience, yet the penetration of agricultural micro insurance remains low compared to other types of micro insurance products. A 2012 study conducted by Making Finance Work For Africa and Munich Re on the landscape of micro insurance in Africa found that agricultural insurance remains at 0.2 million relative to the 39 million in life cover. Africa constitutes the majority of the global population of 1.2 billion people living on less than a dollar a day, of which nearly 75% depend on agricultural activities for their livelihood [?].

Low-income farmers are exposed to both idiosyncratic and aggregate risk, the most common of which is weather risk. Idiosyncratic, non-systematic or individual risk is uncorrelated within a group of farmers. In contrast, aggregate or systematic risks for example floods and droughts, are shocks that affect many people at the same time [?]. Smallholder farmers in developing countries largely depend on rain-fed agriculture and are therefore highly susceptible to weather risk [?]. There is extensive empirical evidence of the adverse impact of downside weather risk on the income earning potential and welfare of households that depend on rain-fed agriculture for their livelihoods. In India a study finds that children from agrarian households are often taken out of school in response to adverse income shocks [?]. A study in Ethiopia finds significant reduction in BMI of children due to rainfall shocks [?]. Another study finds that a 10% lower rainfall about 4–5 years earlier had an impact of one percentage point on current household consumption growth rates [?]. Furthermore, in a survey conducted in 2009 in Ethiopia, 44% of farmers reported serious losses in wealth and consumption due to drought in the last 4 years [?].

Binswanger and Rosenzweig [1993] find that uninsured weather risks are a significant cause of lower farming efficiency and lower average incomes. They further find that a one-standard-deviation decrease in weather risk (measured by the standard deviation of the timing of the rainy season) raises average profits by up to 35 percent among farmers in the lowest wealth quartile [?]. Thus agricultural insurance that reduces weather risk has the potential to improve household welfare for many of the world's poor rural households [???].

Insurance is a financial arrangement widely used in developed countries, it can be a mechanism to protect people against downside risks (or shocks), which include accidents, illness, death in the family, natural disasters and weather risk in agriculture. A loss is covered in exchange for insurance premium payments tailored to the needs, income and level of risk of the insured [?]. Micro insurance is aimed primarily at the developing world's low-income market, people thought to live on less than 2 dollars per day. Life, health and agriculture are the common perils covered. Premiums payments are set proportionate to the likelihood and cost of the risk involved. Premium are typically small and paid frequently, which is suitable for the paying capacity of these communities. Policies are usually simplified and delivered through micro finance institutions rather than the insurance private sector [?].

Many low-income households and especially farmers do have mechanisms to cope with ex-post and ex-ante risk. Ex-ante strategies involve actions that prevent the risk or reduce the probability of the risk occurring. Ex-post strategies are actions that reduce the impact of the risk after it has occurred. Ex-post, consumption is smoothed by reducing expenditure on basics, including health and education [?] and by selling productive assets that have taken years to accumulate and will be hard to replace [?]. However, these ex-post risk management strategies further reduce household welfare. Ex-ante, exposed farmers may smooth income by investing in low-risk, low-return crops and production technologies [??]. However, these ex-ante risk management strategies limit farmers' ability to develop out of poverty and overall are not welfare improving.

An additional risk management mechanism smallholder farmers depend on is their social network. There exists a large body of literature on informal risk sharing in largely rural, low-income households [????????]. Informal risk networks can be described as networks of family, friends and neighbours who engage in reciprocal transfers in cash or kind through institutions that are not regulated by the financial services regulatory body of a country. Notwithstanding that the method of payment or transfer may be formal in that the group may collect premiums regularly and provide a payout to affected members in a coordinated method. However, many studies find that the ability of a social network to fully insure against covariate risks is limited. Without financial resilience, incomplete insurance for smallholder farmers who rely on rain-fed agriculture threatens the current and future welfare of their households [?].

In rural Ethiopia, smallholder farmers voluntarily associate in informal risk sharing networks known as iddirs. These are indigenous, egalitarian and transparent institutions that developed as a mechanism to cope with low frequency, high cost shocks, particularly death. Hence iddirs developed as funeral insurance provision groups. From funeral cover, iddirs evolved to offering other forms of insurance as well as credit to their members [?]. Since 1966, iddirs have been required by law to register as associations [?]. To date, iddirs have stable and clearly defined membership, rules, procedures and payout schedules, including joining fees and fines for late payment of one's contribution. They are long-standing institutions with an average duration of 18.5 years and an average of 85 members who democratically elect their governing bodies. Iddirs can be classified into 3 distinct types based on the extent to which contributions depend on the occurrence of a shock. Type 1 iddirs contribute only upon the occurrence of a shock, type 2 contribute upon entry and the occurrence of a shock, while type 3 contribute only by regular monthly premiums. These contribution methods in turn map onto ex-ante versus ex-post transfers, the first type uses ex-post transfers, the second uses a mixture while the last type of iddir uses only ex-ante transfers [?].

The literature on informal risk sharing shows that informal risk-sharing exists but that it provides incomplete insurance, particularly against aggregate shocks. Townsend [1994] finds that in south Indian villages, although household income varies, consumption is largely smooth. However, the landless and smallholder farmers are more vulnerable to idiosyncratic shocks. Similarly, Ravallion and Dearden [1988] in their study of the moral economy in Indonesia find evidence of informal risk sharing in the form of voluntary interpersonal transfers of money or goods in both urban and rural areas [?]. Likewise, Rosenzweig [1988] also finds that households are an important risk-mitigating institution through transfers [?]. However, this depends on the household structure and its social ties beyond its immediate village proximity.

Even more telling, Rosenzweig *et al.*, [1993] argue that rural agents are successful in insuring against all non-covariant or idiosyncratic risk. However, this successful insurance provision is not village-wide [?]. Fafchamps and Lund [2003] find that mutual insurance, in the form of gifts and flexible, zero-interest informal loans take place primarily through networks of friends and relatives. They argue that the fact that households receive gifts and loans from network members is not, in itself, evidence of inefficient risk-sharing: Efficient risk sharing can still be achieved through social networks provided exchange is frictionless and individual networks overlap so that no villagers are excluded. Nonetheless it is important to note that certain shocks, like death and loss of employment of household head are better insured than others like sickness and loss of employment of dependents [?].

Despite the fact that informal risk sharing (informal insurance) provides incomplete cover and is weak against repeated shocks [??] , studies on micro-insurance demand in rural settings have found that demand for formal insurance is very low. Previous studies have found 16% demand for health micro insurance in Kenya [?] and 20.4% in Nicaragua [?]. A study in India found a 25% demand and that lack of trust and credit constraints limit demand [?]. Even when the insurance premium is subsidised below actuarially fair rates, demand for agricultural insurance in developing countries is low relative to demand in developed countries [?]. Moreover, in developing countries, demand for agricultural micro insurance is lower than demand for other micro insurance products like life and funeral cover [?], with studies reporting a 11.4% demand for rainfall insurance in India [?] and less than 20% take-up rate of agriculture index insurance in a randomised control trial in Ethiopia [?].

There are numerous possible explanations for the low insurance demand. A review of studies on micro insurance demand in the last 13 years identifies 12 key factors that determine demand for micro insurance [?]. These can be grouped into economic factors (price and wealth), social and cultural factors (risk aversion, nonperformance risk, trust, peer effects, religion, financial literacy), structural factors (informal risk sharing, quality of service and risk exposure), and personal and demographic factors (age and gender) [??]. However, among the contending explanations for low insurance take-up, the interaction between informal risk sharing and formal insurance remains under investigated.

Informal risk sharing is an important structural factor which interconnects social and cultural factors like trust and peer effects to economic factors like the quality of service provided by the insurance industry and non-performance risk. Generally, non-performance risk is the possibility that an insurer may not fulfil its obligation to pay an insured policyholder upon a claim. This may arise from contract exclusions, or insurer bankruptcy. Specifically to index insurance, non-perfomance risk arises as basis risk due to a lack of adequate historical weather data or infrastructure for accurate measurement of weather. Contract non-performance in index insurance through basis risk has been seen to limit its take-up. In rural India, take-up of rainfall index is found to increase with wealth and participation in village networks, but decreases with basis risk [?].

Social transfers in pre-existing informal risk-sharing networks provide informal insurance that may condition the demand for formal insurance. Hence, demand for formal insurance has been shown to depend on a household's informal risk-sharing arrangement [???]. While others find that participation in village networks increases take-up, in other words, informal risk sharing crowds in formal insurance [?], others have found results that contradict these findings. An empirical study in the Philippines finds that formal insurance crowds out informal insurance and reduces solidarity in irreversible ways, such that the withdrawal of formal insurance leaves the farmers worse off [?]. However, unlike our study of both indemnity and index insurance, the variations of insurance in their study are both indemnity in nature. Relatedly, a laboratory experiment finds that formal insurance crowds out social transfers in informal risk-sharing networks [?]. Therefore, informal risk sharing mechanisms can either positively or a negatively affect micro insurance demand.

A contemporaneous paper to this study finds that support to iddirs, to strengthen the extent to which they were able to insure members against idiosyncratic shocks increases demand for index insurance [?]. However, the existent literature is inconclusive due to the evidence available to challenge these findings. Moreover, there also exist theoretical findings that the "low" observed demand for weather index insurance by poor farmers may be rational, or even too high [??].

A recent study examines the impact of index insurance and local risk-sharing on welfare, and the complementarity between formal and informal insurance in Ethiopia. The study theoretically shows that the presence of basis risk makes index insurance a complement to informal risk sharing when informal insurance covers idiosyncratic risk and basis risk is uncorrelated among its members [?]. This finding suggests that take-up of index insurance would not result in discontinuity of transfers in informal risk sharing networks. Instead, the informal risk sharing networks would respond to the basis risk when a member takes up index insurance by offering transfers to cover the loss.

Based on existent iddirs in Ethiopia, Dercon *et al.*, [2014] empirically find that groups whose leadership was trained on the benefits of sharing insurance had substantially higher take-up rates than iddirs whose leadership was trained on the individual benefits of insurance. The analysis is inconclusive on whether this means demand for index insurance is higher among those inclined to share risk or insurance marketing that provides messages on risk sharing increases demand. If the former is the case, then assuming that the same iddir pair-types in our study are inclined to share risk, one would expect that the take-up of index insurance would not result in transfers being reduced when there is basis risk.

Another theoretical study, which addresses a different but related question on whether selling insurance to groups would increase demand identifies two characteristics that suppress demand for insurance. When members belong to a group with common interests but the decision to purchase insurance is made individually, take-up can be low due to free-riding and coordination problems. Free-riding arises because one individual's insurance decision creates positive externalities for the other group members, such that they value the insurance less. Coordination problems arise because an individual's insurance can be either positively or negatively valued depending on the insurance decisions of the other group members. This may create a coordination problem whereby not all the members of the entire group decide to buy insurance, an outcome that is Pareto dominant [**?**].

The concept of group insurance in de Janvry *et al.* [2014] means that a group jointly purchases the policy and by default shares the payout triggered by a common shock. On the other hand, Dercon *et al.* [2014] conceptualize group insurance largely in terms of the sharing of a payout corresponding to each individual's loss. The group does not purchase the insurance policy jointly and a common shock triggers a payout. However, members face different levels of idiosyncratic loss and are able to share their individual payouts among the worst affected group members. This notion in Dercon *et al.* [2014] is based on the empirical finding that although groups in their study were permitted joint insurance purchase, only 6% reported sharing policies. Thus here the rationality of group insurance is that informal risk sharing allows for loss adjustment among group members therefore reducing the impact of basis risk. In contrast, de Janvry *et al.* [2014] do not explicitly consider informal risk sharing. However, because their theoretical model uses indirect utilities defined over both individual and aggregate consumption, it accounts for interactions among group members to explain low demand for and crowding out of informal risk sharing. Hence it implicitly accounts for informal risk sharing.

Another related recent study [?] theoretically and empirically examines how the interaction of informal risk-sharing and basis risk impacts the demand for index insurance. Although their empirical work is based in rural India, it is comparable to our Ethiopian study due to the existence of sub-caste groups known as jatis. These are spatially dispersed networks across villages and districts in Indian rural communities, which potentially insure both aggregate and idiosyncratic risks. In their experimental design, they randomize the offer and the price of the index insurance policy as well as the placement of automatic rainfall stations in a subset of the sampled villages. The estimated relationship between the rainfall recorded at the stations and individual farmer's per-acre output provides evidence of basis risk. Whereas their study uses distance from the weather station to a farm as a proxy for basis risk and varies the distance between the farm and the rainfall station as a way of randomize basis risk, we randomize basis risk by the drawing of a token to determine an aggregate shock and the rolling of a die to determine an idiosyncratic shock.

In their study, Mobarak and Mark Rosenzweig [2012] combine Arnott and Stiglitz [1991]'s theoretical model with Clarke [2011]'s theory of rational demand for index insurance and show that, independent of informal insurance against idiosyncratic losses, in the absence of basis risk, farmers choose full, actuarially fair index insurance (which is essentially similar to indemnity insurance). However, if index insurance carries basis risk and informal risk sharing networks protect against idiosyncratic risk then informal risk sharing increases the demand for formal index insurance. Hence the negative effects of basis risk on the demand for index insurance are reduced for those who informally share risk. Thus they show that basis risk creates complementarity between informal and index insurance. Therefore, informal insurance is both a complement and a substitute to formal index insurance depending on the level of basis risk and the nature of the informal risk-sharing arrangement.

The practical implication in Dercon *et al.* [2014] and Mobarak and Mark Rosenzweig [2012] is that groups that are better able to insure individual losses may have a higher demand for index insurance. In other words, members of a network who take up index insurance can rely on their partners to provide cover for idiosyncratic basis risk. Hence in our study, whereby a donor farmer is exempt from losses, one would expect that transfers would increase when there is basis risk compared to the baseline of no index insurance. However, unlike Dercon *et al.* [2014], Mobarak and Rosenzweig [2012] do not offer a conceptual form of group index insurance in terms of making intra-group payments and transfers. So we exercise caution when comparing our results to their findings.

The roots of the insurance market in industrially developed societies can be traced to intergenerational informal risk sharing that was prevalent in agrarian societies. Informal risk-sharing mechanisms in today's emerging economies are not dissimilar. The development of insurance markets formalized the informal risk sharing systems into mutuals and cooperatives, which in turn evolved into national social security. Social security for health and disability has been shown to have both positive and negative impacts on the demand for formal insurance in the developed economies [?]. Similarly, in emerging economies scholars find mixed results on the impact of informal risk sharing on the demand for micro insurance [?].

We find that among the contending explanations for low insurance take-up, the interaction between existing informal risk sharing and formal insurance remains under investigated. Furthermore, development practitioners are increasingly seeking to design and implement formal insurance programmes regardless of whether such informal risk-sharing groups exist or not. The question of the possible effects of formal micro insurance on informal mechanisms in prescriptively formed groups vis-a-vis pre-existing groups is largely overlooked. An understanding of the context in which informal risk sharing and formal insurance are mutually beneficial is critical not only to making valuable micro insurance interventions but to developing a sustainable micro insurance market.

1.3 Predictions and motives for transfers

In this section we draw expectations based on the research design and informed by the literature. We further link the change in transfers to the likely underlying motives.

This thesis primarily investigates *how* transfers change in response to the unobserved behaviour (decisions on insurance) of other members in a network where **group norms of sharing exist**. This is critical to understanding how agricultural insurance interventions impact informal risk-sharing. Interventions range from those that simply make insurance available to those that enforce take-up. Thus in order to understand how transfers change when index insurance is introduced but farmers cannot observe each other's insurance decisions, we examine three possibilities; namely, insurance availability, insurance take-up and insurance rejection.

Furthermore in the absence of pre-existing informal risk-sharing networks, some interventions prescribe farmers into groups. Therefore, the second contribution of this study is on how transfers change in response to unobserved behaviour of other members in prescribed groups, where **group norms of sharing do not exist**.

Related to the question of how transfers change is *why* transfers change. Theory on underlying motives for sharing in rural village networks identifies motivations for sharing as either preference-dependent or incentive-driven [**?**]. A sharer's preferences determine a *baseline (B)* amount that an individual shares as well as an amount directed at improving a particular recipient's welfare, which is called *directed altruism (A).* The level of sharing may also be set in response to socio-economic incentives, namely *sanctions* and *reciprocity (R)* . Ligon and Schechter [2012] identify a sharer's avoidance of social sanctions and the sharer's need for reciprocity as such motives. In contrast, our study is concerned with how a farmer responds to another's insurance decision; therefore we identify the sharer's agency in *punishing (P)* or *rewarding (R)* rather than his or her avoidance of being sanctioned as one potential motive for transfer.

Reciprocity is assumed to lead to complementarity between informal risk sharing that covers idiosyncratic risk and individual index insurance when it carries basis risk [**???**], hence scholars have used dynamic games to investigate this [**???**]. We do not assume reciprocity. In fact we prevent reciprocity: Firstly, by making the recipient of a sharer's transfer anonymous and secondly by playing a one-shot game. Because the recipient has no opportunity to reciprocate the transfer, a transfer cannot be motivated by reciprocity. This leaves only the motive of punishing or rewarding the recipient for their insurance decision.

We chose to focus on factors that depend on socio-economic incentives because the fundamental aim of economic development interventions is to manipulate economic incentives rather than alter preferences. Although our focus is on incentive-driven motives, it is important to account for the fact that transfers are preference-dependent. Hence we elicited a baseline transfer level without insurance for a pre-existing network (PT_S) and a prescribed network (PT_D), in Trt1, which is the control. Next, we introduced insurance and recorded the amount a farmer is willing to transfer when insurance is available. This is the analytical variable T_{ij}^{avail} in Trt2 and Trt3, the response variable under insurance availability. Thereafter, we recorded the amount a farmer was willing to transfer on condition that the recipient took up the insurance offered. This is the analytical variable $T_{ij}^{take-up}$ in Trt2 and Trt3 and the response variable under insurance take-up. Finally, we recorded the amount a farmer is willing to transfer on condition that the recipient rejected the insurance offered. This is T_{ij}^{reject} in Trt2 and Trt3 and the response variable under insurance rejection.

Table 1.1 summarizes a recipient's net pay-off under each insurance condition (availability, take-up or rejection), for both indemnity and index insurance in the event that an idiosyncratic shock is realized. An idiosyncratic shock triggers an indemnity insurance payout, therefore the net pay-off of "Y Take-up" is equal to that of "B Take-up" in row 2 corresponding to indemnity insurance. In contrast, an aggregate shock triggers an index insurance payout. Hence the net pay-off of "Y Take-up" in row 3 corresponding to index insurance is 82 ETB, while under the "B Take-up" condition a recipient does not receive an index payout and therefore suffers basis risk. An aggregate shock does not trigger an indemnity insurance pay-out. However, it does alter the probability that an idiosyncratic shock materializes.

Note that table 1.1 does not show the net pay-off in the event that a recipient does not suffer an idiosyncratic loss. This will be shown later, in Chapter 2, where we use tree diagrams to show precisely how the game unfolds. The diagrams show the likelihood of an aggregate shock, how the aggregate shock determines the probability of an idiosyncratic shock and the recipient's net pay-off at the end of the branch, with or without an idiosyncratic shock. The diagrams also indicate the states in which a donor may make a transfer.

A donor incurs neither idiosyncratic nor aggregate shock and is only asked to make a transfer if a recipient suffers an idiosyncratic loss. If a recipient does not incur an idiosyncratic loss, regardless of the aggregate shock, the donor is not asked to make a transfer. Hence, table 1.1 is a summary of the information that the donor has at the point of deciding how much to transfer. It therefore forms the basis of our expectations of the donor's transfer behaviour.

Without an aggregate shock, we expect the transfers to increase when index insurance is available and when it is taken up because the net pay-off, which is either 10 ETB or 28 ETB is potentially worse than a certain net pay-off of 28 ETB in the control. However, when there is an aggregate shock, we expect a decrease in transfers when index insurance

	No Aggregate shock (B)			Aggregate shock (Y)		
Treatment	B Avail	B Take-up	B Reject	Y Avail	Y Take-up	Y Reject
Trt1 (no insurance)	28	28	28	28	28	28
Trt2 (indemnity)	28 or 70	70	28	28 or 70	70	28
Trt3 (index)	10 or 28	10	28	28 or 82	82	28

Table 1.1: Net pay-off (in ETB) if a recipient has an idiosyncratic loss and insurance is available, taken up or rejected, either without an aggregate shock (B) or with an aggregate shock (Y).

is taken up as the net-payoff is 82 ETB. On the other hand, regardless of the aggregate shock, we expect transfers to decrease when indemnity insurance is taken up due to the overall better net pay-off of 70 ETB. When formal insurance (both indemnity and index) is rejected, the net pay-off is equal in Trt1, Trt2 and Tr3, neither increase nor decrease is warranted. We therefore expect transfers to be equal.

Deviations from the aforementioned inferences may suggest that transfers are not only based on the net pay-off that a recipient has. Hence, donors (narrowly defined as farmers who decide how much to transfer to another) are not simply making needs-based transfers to their partners. Instead, donors may also be reacting to recipients' purchase decisions by either punishing or rewarding them.

If, in any condition, we observe an increase in transfers compared to Trt1 when the net pay-off is higher than that of Trt1, then that is an unwarranted increase in transfers. We therefore posit reward to be the underlying motive. Alternatively, if we observe a decrease in transfers compared to Trt1 when the net pay-off is less than that of Trt1, then that is an unwarranted decrease in transfers. We thus posit punishment to be the driving motive.

When there is no aggregate shock and index insurance is taken up, there is basis risk. Therefore, an increase in transfers compared to the control is warranted. In which case, index insurance and transfers in informal risk sharing are supplementary. In contrast, if transfers decrease compared to the control, then it constitutes an unwarranted decrease in transfers. This may suggest a punishment motive, which may be reasonable if a donor views index insurance as risky and the donor's decision to take it up as unwise. Note that supplementarity in this study is defined as either a warranted increase in the level of transfers (or odds of a transfer).

Chapter 2

Research design and data

In this chapter, we discuss the sampling design and describe our framed field experiment. The study conducted field experiments in Tigray province in Ethiopia over a period of 5 weeks, covering 1536 low-income farmers, most of whom were selected from a sample of 16 pre-existing informal risk-sharing groups called iddir. Due to randomly missing data, our sample size reduced to 1527 farmers. The research design is from a broader study that examines risk preferences and motives for transfers in risk sharing groups. It examines both preference-dependent factors and factors due to socio-economic incentives extensively. It predicts that the complex interaction of these factors affects the extent of complementarity or substitution of formal insurance in informal risk sharing networks [?]. Although the experiments were conducted jointly with Morsink [communication], this study addresses a different set of questions and conducts a separate analysis.

For the informal risk-sharing experiment, a donor farmer was paired with a recipient farmer. The recipient in each pair was offered a chance to purchase either index or indemnity insurance. In order to compare the effects in pre-existing groups to prescribed groups the sample is stratified into two Pair-Types; pairs of farmers from the same iddir (PT_S) and pairs of farmers from different iddir (PT_D). The farmers were made to understand that their final payoff at the end of the game would be determined by their choices in the game. This ensured that farmers' decisions were driven by economic motives rather than social pressure, impressing their local district administrator or conveying any image of himself or herself to the experimenter, Section 2.2 elaborates on this.

2.1 Mixed randomised design

We conducted a lab-in-the-field experiment in the form of a one-shot, two-person, informal risk-sharing game. Two-person means that the game is played in pairs. One-shot means that we restrict the transfers to one season and one direction, hence a transfer can only flow from a donor farmer to a recipient farmer. This prevents reciprocity and enables us to to understand how one farmer's insurance purchase decision affects the transfer decision of another. The objective of the game was to understand how one farmer's decision to take up index insurance affects the willingness of his or her partner to make a transfer in case of a loss. In each pair, the recipient farmer received an offer of either index or indemnity insurance and was subject to a shock. On the other hand, the donor farmer did not experience a loss and was asked to make a decision on how much he or she would transfer if his or her partner suffered a loss. We also conducted an observational study in the form of a survey, which collected data at the iddir and farmer levels [**?**].

Statistical theory distinguishes between *hierarchical design*, which is described as an observational study in which the categorical variables are associated with the participants before the study (for example, gender, iddir, village, marital status) and *split-plot design*, which is an experimental design whereby treatments are randomly allocated to subjects and their responses recorded [?]. Our study employs a hierarchical observational design and a split-plot experimental design and is best described by both and not one or the other.

2.1.1 Hierarchical sampling design

The study used a 3-stage sampling design with farmers selected from iddirs, iddirs within villages and villages within administrative units in Tigray province. In stage one we select 3 administrative units, within which we randomly select a total of 8 villages, 3 from each of 2 administrative units and 2 from one administrative unit. In stage two we randomly select 2 iddirs from each of the villages which gives a total of 16 iddirs. In stage three we then randomly select 72 farmers from each iddir and 24 farmers from a different iddir but within the same village as the 72 previously selected farmers. Thus the farmers are not independently sampled since the chance that a farmer will be selected depends on whether the iddir that they are a member of is selected or not in the first stage. This is presented in Fig. 2.1.

Our hierarchical structure is such that we selected farmers within iddirs, iddirs within villages and villages within administrative units. The three administrative units were selected because of their proximity (45 minutes – 180 minutes drive) to Mekelle, the capital city of Tigray province, as well as their accessibility. Villages were randomly selected from a list of villages in each administrative region that had a minimum of six iddir per village. The number of farmers in all 16 iddirs ranged from 100 - 200. Overall, we had 1081 farmers selected from the sampled iddirs and 353 farmers were selected from

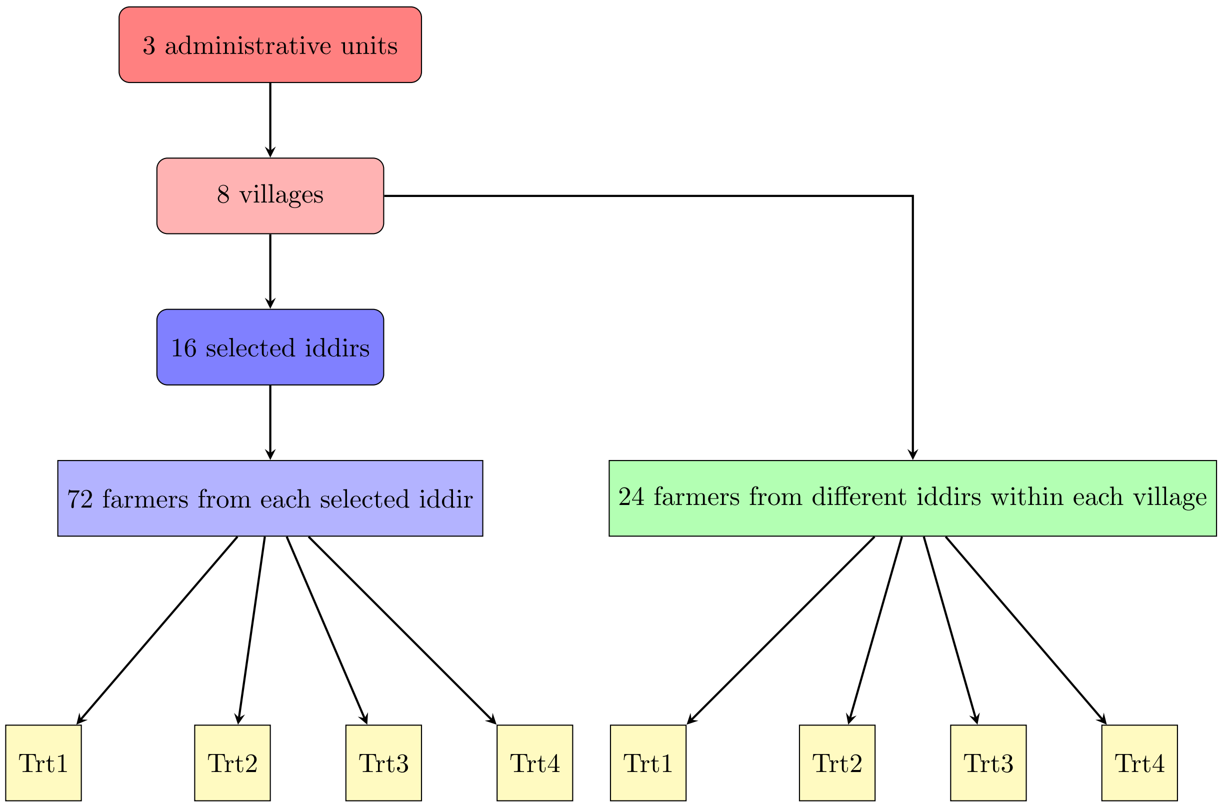

Figure 2.1: Hierarchical sampling design

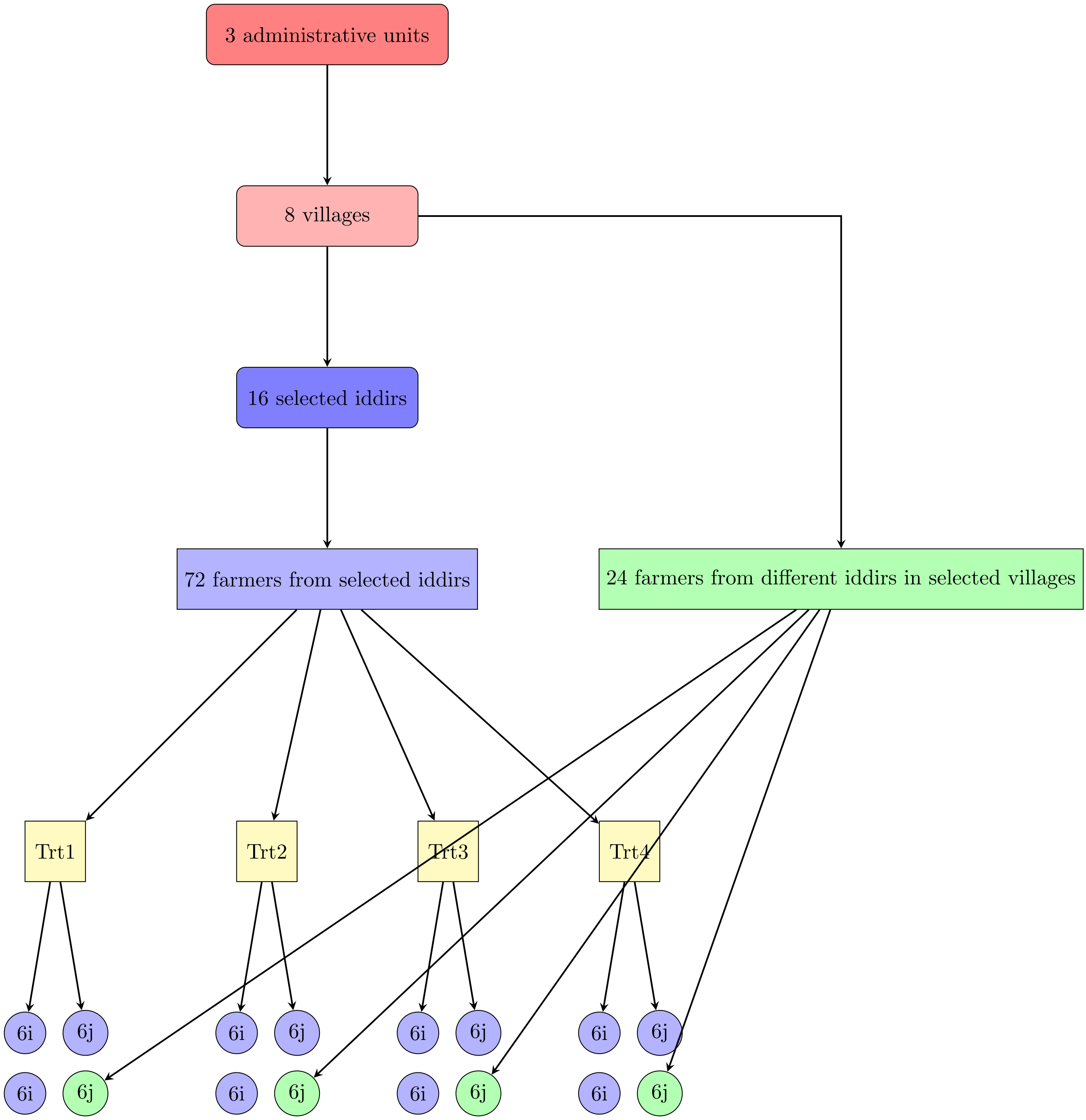

Figure 2.2: Hierarchical sampling and pairing design

the sampled villages but not iddirs; these were paired with a third of the farmers from the sampled iddirs while the remaining two thirds were paired with each other. Fig. 2.2 extends on the sampling design to illustrate the pairing and the levels in our hierarchical sampling, indicating the stratification in the sample to create 6 PT_S (i and j both purple) and 6 PT_D (i in purple and j in green). We describe the experimental design in the next section.

2.1.2 Experimental split-plot design

Experimentally, we recorded transfers from 6 donors in each of two pair-types (PT_D and PT_S) within each of the four treatments in each of the 16 iddirs, which is illustrated in Fig. 2.3 and presented more compactly in Fig. 2.4.

The informal risk-sharing game has 4 treatments, which all iddirs receive. We randomly assign each of the 72 farmers to one of the treatments, so that each treatment has 18 farmers from the same iddir. Similarly we randomly allocate each of the 24 different-iddir farmers to either of the 4 treatments. The farmers are then randomly paired, resulting in 6 pairs in which farmers belong to the same iddir (PT_S) and 6 pairs in which farmers are from different iddirs (PT_D). Therefore because the 6 observations in each treatment from the same iddir:pair-type combination are not independent, the data has *pseudo-replication*, which violates the linear modelling assumption of identically and independently distributed errors. [**?**].

For PT_S, we randomly assign one farmer the role of recipient and another donor. However, for PT_D the farmers from the sampled iddirs in stage two are assigned as donors

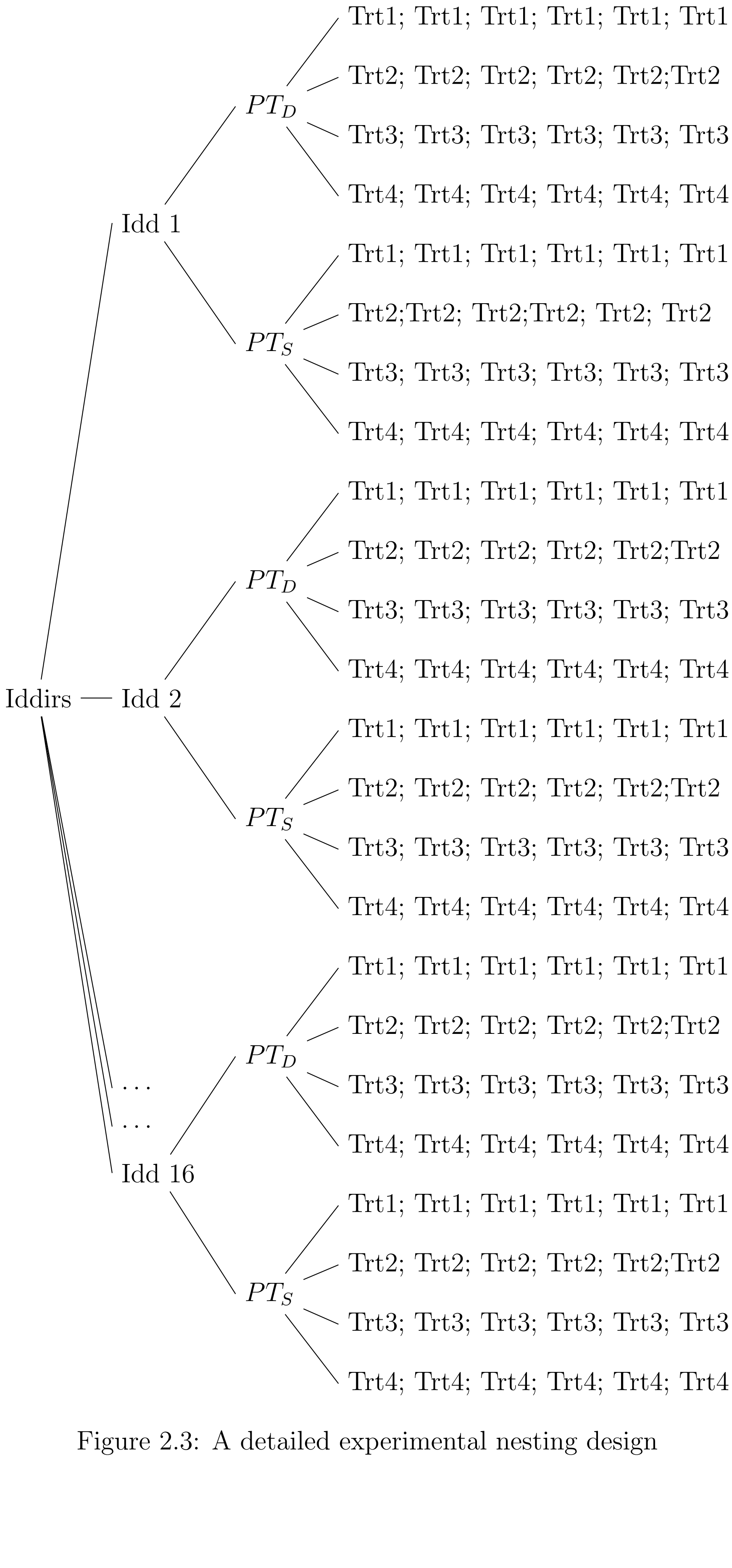

Figure 2.3: A detailed experimental nesting design

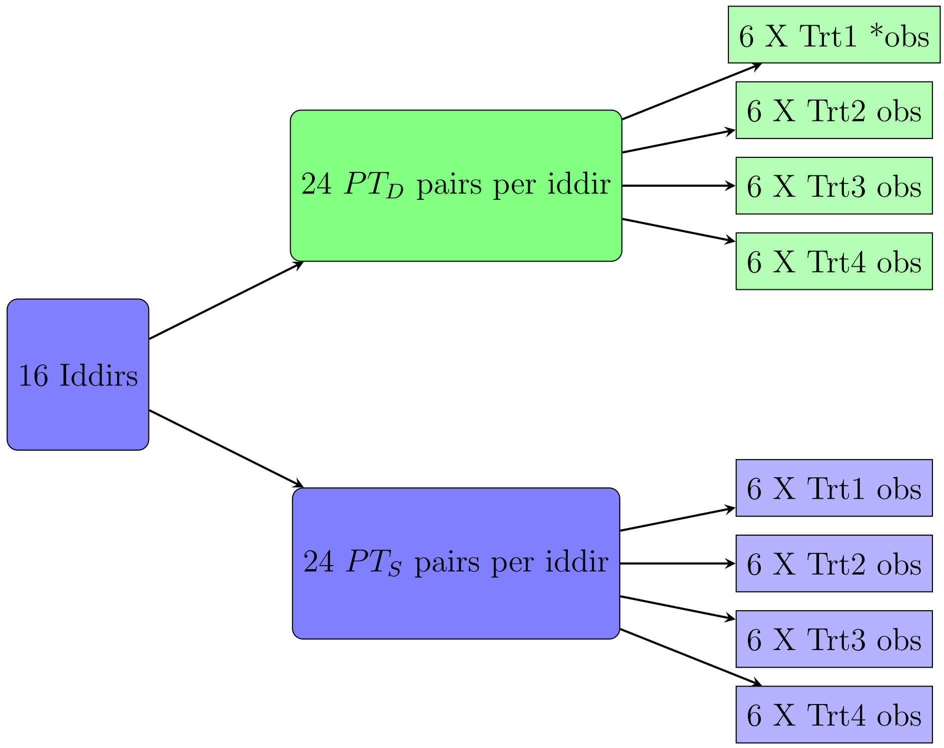

Figure 2.4: A compact experimental nesting design. *obs stands for observations

and by default, farmers from a different iddir are assigned as recipients. Therefore, of the 18 farmers from the selected iddir, 12 are assigned as donors and 6 assigned as recipients. This is visually demonstrated in Fig. 2.4, which also depicts the 6 farmer pairs nested in Pair-Types, which in turn are nested in treatments, in turn nested in iddirs in our experimental design.

2.1.3 Imperfect Hierarchies

Due to the nature of framed field experiment, which we elaborate on in Section2.2, the use of ten enumerators adds a further complication to the design, namely *imperfect hierarchy*, which is a situation whereby lower level units are nested in more than 1 higher level unit. In this case each farmer belongs uniquely to an iddir (higher level unit) and uniquely to an enumerator (a second higher level unit). Enumerators are not nested within iddirs, but crossed with the iddir factor because each enumerator is the same across each iddir. Fig. 2.5 illustrates the full experimental design.

2.1.4 Randomisation

We randomise the selection of iddirs, the selection of farmers within iddirs, the allocation of farmers into treatment and pair-types and the assignment of each farmer as either a donor or a recipient. Randomisation helps address some of the problems endemic to social science research, particularly selection bias, which occurs when subjects are selected based on characteristics that may also affect the outcomes. Such a situation makes it is difficult to disentangle the treatment effect from pre-existing differences [**?**].

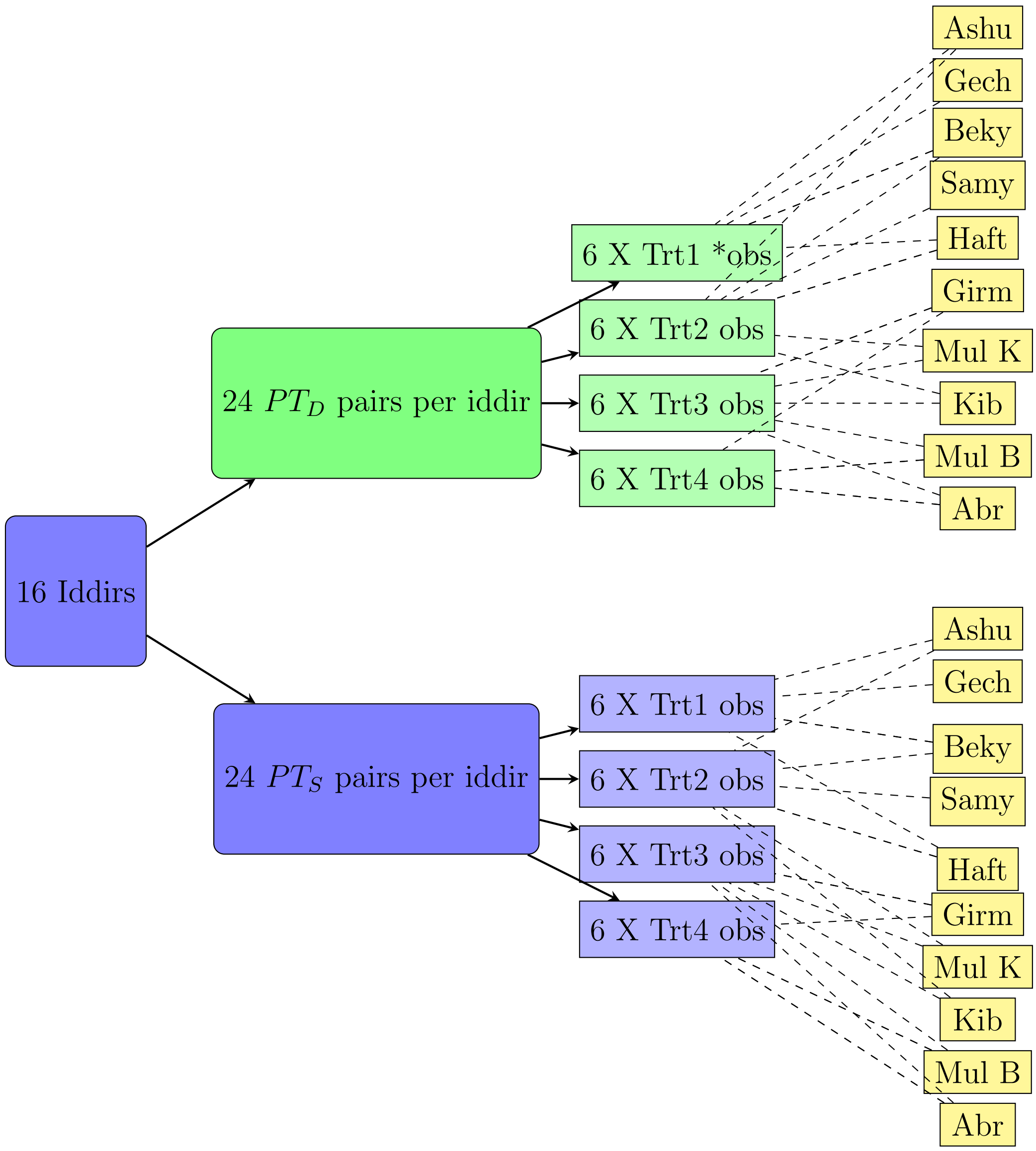

Figure 2.5: Full experimental design; Treatment pair-type nested in iddirs and crossed with an enumerator. Note: *obs stands for observations

Randomisation is necessary because our aim is to make causal inferences about the processes captured in the data. As Crawley argues, understanding causality necessarily requires answering a counterfactual [?]. In our case, the question;

"What is the causal effect of the availability of index insurance on transfers?"

necessarily involves answering the question,

"What do the transfers look like, ceteris paribus, in the absence of index insurance?"

Comparing the responses of one individual before and after the introduction of insurance may be one way to investigate this. However, this creates a learning effect, which is highly dependent on the respondents selected. Instead randomisation and comparing a treatment group to a control group such that pre-existing differences are balanced in each group prevents both selection bias and a learning effect.

Randomisation also helps minimise systematic errors whereby the experimental conditions are consistently different for the treatment and control group due to other independent variables, termed confounding variables. Finally, although randomisation does not eliminate publication bias, it helps minimise it. Statistically insignificant or unexpected results are less likely to be dismissed as erroneous. The results can still provide useful information and may be published. The experimental treatments are:

- Trt1: all recipients are offered no insurance
- Trt2: all recipients are offered indemnity insurance
- Trt3: all recipients are offered index insurance
- Trt4: a third of the recipients are offered no insurance, a third are offered indemnity insurance, and a third are offered index insurance (per iddir)

In the next section, we elaborate on the risk-sharing game and contextualize our study in the wider field of framed field experimental research designs with economic incentives.

2.2 Field experiment with economic incentives

The aim of field experiments in social science is to observe subjects in a controlled setting without imposing controls that the subjects may perceive of as being unnatural or deceptive. The use of field experiments in development economics has seen a surge in the last decade due to their power to incorporate context which is fundamental to human behaviour and critical for understanding causality. Notwithstanding, there are limits to replicability when using a frame-field study. Replicating the research design using a different sample or retesting the hypotheses of the original study using a different research design could lead to different results [**???**].

Framed field experiments draw subjects from a naturally occurring population and incorporate elements of the naturally occurring environment, like the commodity, task, stakes, and experience or information of the subjects. In our case, naturally occurring and identifiable informal risk sharing networks (iddir) are the object of our study, hence we sample farmers who are members of an iddir. Exposure to aggregate and idiosyncratic agricultural shocks and familiarity with sharing norms was important context-specific experience hence we select only farmers. In addition to providing economic incentives, which ensured that the stakes were high, the subject matter under study was meaningful to the farmers. This gives us confidence that the farmers participated with understanding.

In addition, the meaningfulness of the subject enabled high levels of understanding among our sample of farmers. We find that although farmers had attained low levels of formal education, they understood the aim of the study and the design of the game. In each session, after explaining the game, we asked participants a series of questions to test their levels of understanding. The results of these tests for the respective treatments are discussed in the subsequent sections.

Our study involves randomisation of a naturally-occurring population of smallholder farmers in their natural environment, with the research subjects being aware that they were part of an experiment. The subjects' awareness of being monitored, recorded, and scrutinised may have introduced randomisation bias and compromised generalisability. However, framed field experiments counter this by allowing the researcher to work with subjects in their natural environment. Psychology literature also documents the disempowering impact of being an experimental subject, vis-à-vis the power of the experimenter and the experimental situation [?] and [?]. This may have a negative effect on the quality of the data. However, we counter this problem by providing economic incentives. The income from the experiment was a substantial proportion of the daily wage of a labourer during this period in rural Tigray.

During the recruitment phase, subjects were informed that they were eligible to participate in an experiment in which they would be asked to play two games, a risk sharing game and a risk preference game and to participate in a survey [?]. The risk preference game is beyond the scope of this study, hence we will not elaborate on it any further in this thesis.

Subjects were told that in the risk-sharing game, they would be teamed up with a farmer either from their own or a different iddir and would be asked to make decisions about taking up insurance or making a transfer. Farmers were also informed that they would receive a base-payment of 50 Ethiopian Birr (ETB) irrespective of the outcomes of their own or their partner's decisions in the games. They were also informed that they would be able to win an additional 0–110 ETB depending on their own or their partner's decisions in the games. The additional payment would be based on the outcome in either of the two games, based on a random draw at the end of the session. Subjects were also informed that the total participation time including the two games, the survey and the payment would not be more than three hours. The incentives in the games reflect a daily wage for unskilled labour, which ranges between 50 and 150 ETB during the timing of the experiments [**?**].

Each day of the 5 weeks during which we ran the experiments, we conducted a morning and an afternoon session. In each session we had two classrooms in which a session leader oversaw the experiment. We also randomly allocated enumerators to session leaders. Each session began with an instruction phase during which we explained the objective of the study and the procedure of the game. Participants were informed that there would be two types of farmers in their session, a *donor* and *recipient*, and that they would be randomly assigned to be either one of these in the course of the game. They also learned that $\frac{3}{4}$ of the farmers were members of their own iddir and $\frac{1}{4}$ of the farmers were members of another iddir. Farmers were instructed that, after they had been assigned to be either a donor or recipient farmer, they would be randomly matched with another farmer to form either a PT_D or PT_S pair [**?**].

Prior to conducting the experiment, the session leader in each room explained the risk-sharing game to all the subjects. We showed by examples the drawing of a token and the rolling of a die, dependent on the token drawn. We showed them how these events determined the outcome for the recipient farmer's production. We also illustrated the outcomes of each insurance decision with their corresponding implications on a recipient's net pay-off. Table 1.1 summarizes the net pay-off in the event that an idiosyncratic loss materialized. However, the farmers were shown the net payoff with or without an idiosyncratic loss. The enumerators translated into the local language. Before proceeding to play the game, all subjects were asked a series of questions to test their level of understanding of the game.

2.2.1 The risk-sharing game

A donor and recipient initially have 0 ETB and both can earn an income of 100 ETB from agricultural production. A donor earns 100 ETB from his or her agricultural produce with certainty but a recipient's agricultural production is subject to an idiosyncratic loss (denoted S_r), which can lead to agricultural produce with a value of 28 ETB. If an idiosyncratic shock occurs, S_r takes a value of 0, otherwise $S_r = 1$. The probability of idiosyncratic loss is determined by whether or not an aggregate weather shock (denoted A) is realized. The aggregate weather shock (A) takes a value of 1, which meant good weather, with probability $\frac{3}{4}$ and value of 0, which meant bad weather, with probability $\frac{1}{4}$.

The state of A determines the probability of an idiosyncratic loss, hence the idiosyncratic loss is state-dependent. If $A = 0$ the probability of idiosyncratic loss is $\frac{2}{3}$, which provides an income of 28 ETB while the probability of no loss is $\frac{1}{3}$, which provides an income of 100 ETB. If $A = 1$ the probability of idiosyncratic loss is $\frac{1}{3}$, which provides income of 28 ETB and the probability of no loss is $\frac{2}{3}$, which provides income of 100 ETB.

The game consisted of five steps:

Step 1: The donor receives an income of 100 ETB

Step 2: The realization of an aggregate shock (which is known to the donor)

Step 3: The donor decides how much to share

Step 4: The recipient realizes an idiosyncratic loss

Step 5: The actual transfer from donor to recipient

In step 1 the realisation of income was simulated in the game as the realisation of either a good crop or a bad crop. Farmers were informed that both donor and recipient could have a good crop which would lead to them receiving an income of 100 ETB. Farmers also learned that the donor would have a good crop with certainty and receive 100 ETB but the recipient would be subject to an idiosyncratic loss which could result in a bad crop and would reduce his income to 28 ETB. Participants were informed that the donor would be asked to make a decision to share part of his or her income of 100 ETB with a recipient if the recipient experienced an idiosyncratic loss of 72 ETB resulting in an income of 28 ETB.

In step 2 the realisation of an aggregate shock was simulated by a random draw from an envelope which contained three blue tokens, representing good weather, and one yellow token, representing bad weather. Hence the probability of an aggregate shock (A= 0) was $\frac{1}{4}$. The aggregate shock was determined centrally by the session leader drawing a token in full view of all the donors and recipients in that room. The outcome determined which die the recipient would individually roll in step 4.

In step 3 the donor was asked to make a strategic decision about risk-sharing. Firstly the donor was asked how much he or she wanted to transfer to a recipient partner in the event that the recipient had an idiosyncratic loss. This amount is used as Tij^{avail} in Trt1, Trt2 and Trt3. The donor did not know the recipient's insurance decision, hence in deciding how much to transfer, he or she made a strategic decision based on the net pay-off shown in table 1.1. Columns 1 and 4 show the net pay-off when insurance is available, without an aggregate shock and with an aggregate shock respectively.

In Trt2 and Trt3, after the donor had made a transfer decision under the condition of insurance availability, the donor was asked how much he or she wanted to transfer to a recipient partner in the event that the recipient had an idiosyncratic loss and had taken-up the insurance offered. This amount is used as $Tij^{take-up}$. Again, the donor did not know the recipient's insurance decision. However, the donor knew the net pay-off from insurance take-up. These are given in columns 2 and 5 of table 1.1 for the condition without and with an aggregate shock, respectively. Finally, the donor was asked how much he or she wanted to transfer to a recipient partner in the event that the recipient had an idiosyncratic loss and had rejected the insurance offered. This amount is used as T_{ij}^{reject}. The donor did not know whether the recipient had decided not to take up the insurance offered. However, he or she knew the resultant net pay-offs from rejecting insurance, without and with an aggregate shock, which are respectively given in columns 3 and 6 of table 1.1.

In step 4 we used two dice (red and white) to simulate an idiosyncratic loss conditional on the aggregate shock. A red die was rolled if there was no aggregate shock (A = 1) whilst a white die was rolled if there was an aggregate shock (A = 0). The red die had four blue sides, representing a good crop (100 ETB), and two yellow sides, representing a bad crop (28 ETB). Hence the probability of an idiosyncratic loss when the weather was good was $\frac{1}{3}$. The white die had two blue sides and four yellow sides, thus the probability of an idiosyncratic loss if there was an aggregate shock was $\frac{2}{3}$. The idiosyncratic shocks were determined individually by each recipient rolling a die while an enumerator observed.

In step 5, the outcomes of steps 3 and step 4 are revealed to the recipient and donor, respectively. Depending on the recipient's insurance decision, the relevant amount was then transferred from the donor to the recipient.

Trt1, Trt2 and Trt3 allow for between subject comparison, while Trt4 allows for within-subject comparison. Henceforth we exclude Trt4 as our analysis is focused on between-subject comparisons. The rest of the chapter discusses the details of each of these treatments.

2.2.2 The control, no insurance risk-sharing treatment (Trt1)

The randomly selected recipients individually and secretly rolled a die to determine an idiosyncratic loss while the donor was asked to decide how much he or she wanted to transfer to his or her partner (recipient) if they experienced an idiosyncratic loss.

The probability distributions and possible outcomes for recipients in Trt1 are presented in Fig. 2.6. The conditional probability of idiosyncratic loss is thus $\frac{5}{12}$. A donor's transfer, T_{ij} is contingent on a recipient experiencing an idiosyncratic loss. Hence, T_{ij} is added at

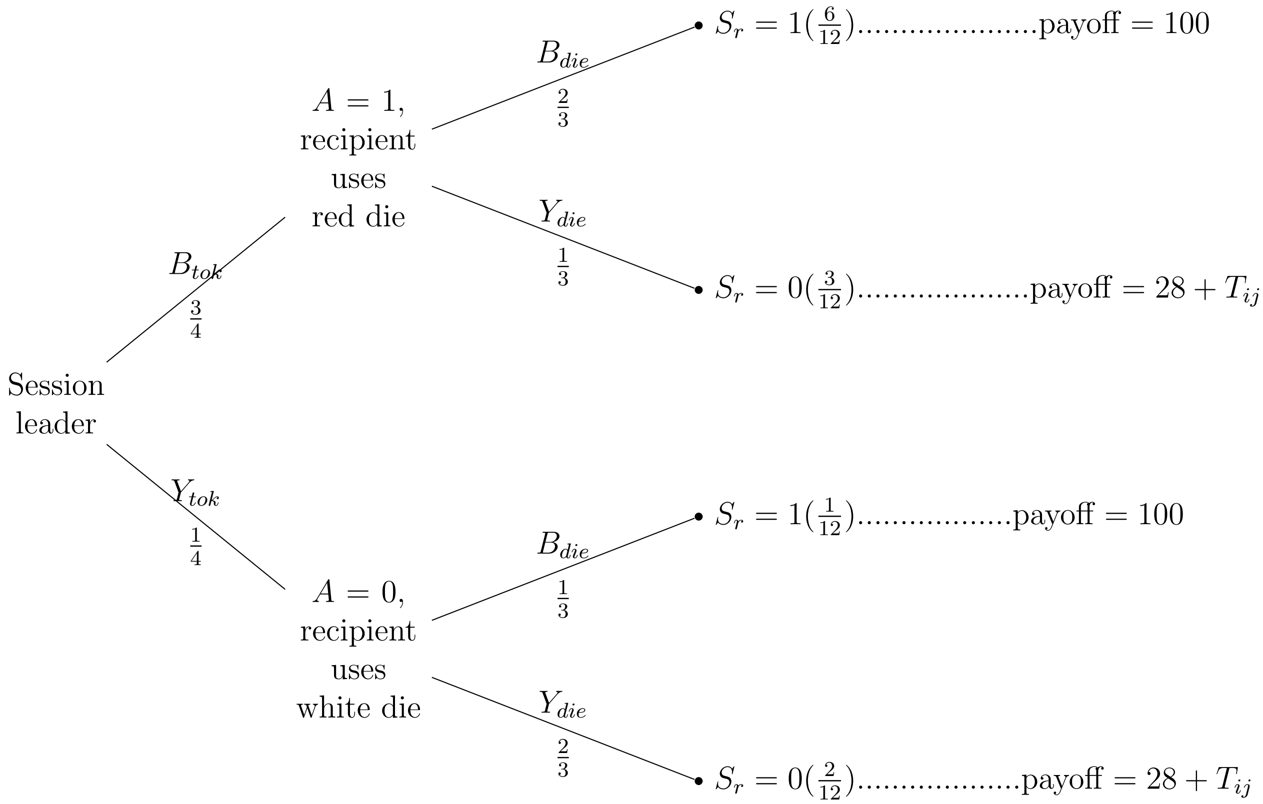

Figure 2.6: Probability distribution and payoff structure in Trt1: state-dependent idiosyncratic loss. Where B_{tok} represents drawing a blue token and Y_{tok} a yellow token and B_{die} represents a blue outcome from the die roll and Y_{die} a yellow outcome.

the end of the branches along which an idiosyncratic loss occurs. The actual transfers were made if a recipient experienced an idiosyncratic loss.

2.2.3 The indemnity insurance risk-sharing treatment (Trt2)

In the indemnity insurance treatment, step 1 has two components. While the donor receives an income, the recipient is offered actuarially fair indemnity insurance with premium m_m and insurance payout P_m. The payoff is chosen to provide full indemnity, implying that $P_m = 72$.

For insurance to be actuarially fair, the insurer should have an expected profit of zero. Suppose the insured pays r ETB for 1 ETB worth of insurance. Thus for x ETB of insurance the premium is rx ETB. The probability that the insured suffers a loss is p. Therefore, the insurer pays out x ETB with probability p and with probability $(1-p)$ the insurer does not make a pay-out. Setting the insurer's expected profit to zero;

$$\begin{aligned} p(rx-x)+(1-p)(rx-0) &= 0 \\ px(r-1)+(1-p)rx &= 0 \\ pr-p+r-pr &= 0 \\ r &= p. \end{aligned} \tag{2.1}$$

Therefore, being actuarially fair, the premium rate must be equal to the probability of a loss.

In the case of indemnity insurance, the probability of a loss is the probability of an idiosyncratic shock, which from the probability distribution of a Trt2 recipient, given in Fig 2.7 is $\frac{3}{12}+\frac{2}{12}=\frac{5}{12}$, thus

$$m_m = \tfrac{5}{12} * P_m = 30\,\text{ETB where } P_m = 72 \text{ ETB, if } S_r = 0.$$

It was explained to the participants that in case they chose to take up insurance the premium would be deducted from their income in the game while if they chose not to take up the insurance the premium would not be deducted. Recipients made insurance decisions privately, such that their decisions were unknown to their donor partners.

After a donor had received the initial income of 100 ETB and a recipient made a decision on the insurance offer, the session leader drew a token centrally, to determine an aggregate shock. Thereafter, the donor was asked to make a strategic decision on how much he or she would like to transfer if a recipient partner suffered an idiosyncratic loss. We further asked how much the donor would share conditional on the recipient's insurance decision, namely under take-up and rejection. These responses are T^{avail}, $T^{take-up}$ and T^{reject} respectively.

The realization of an idiosyncratic shock to the recipient was identical to the control treatment. In the final step of the game, the recipient received an insurance payout if he or she had purchased insurance and experienced a loss. In addition, the recipient also received a transfer based on the strategic decision by the donor in step 3 of the game. The probability distributions and possible outcomes for recipients in Trt2 are presented in Fig. 2.7. A donor's transfer, T_{ij} is contingent on a recipient experiencing an idiosyncratic loss. In addition, a donor made transfer decisions $T^{take-up}$ and T^{reject} conditional on the recipient's insurance decision. These conditional transfers are added to the end of the branches along which an idiosyncratic loss occurs. The actual transfers were made if a recipient experienced an idiosyncratic loss. An amount $T_{ij}^{take-up}$ was transferred if the recipient had taken up insurance in step 1 while an amount T_{ij}^{reject} was transferred in the recipient had rejected the insurance.

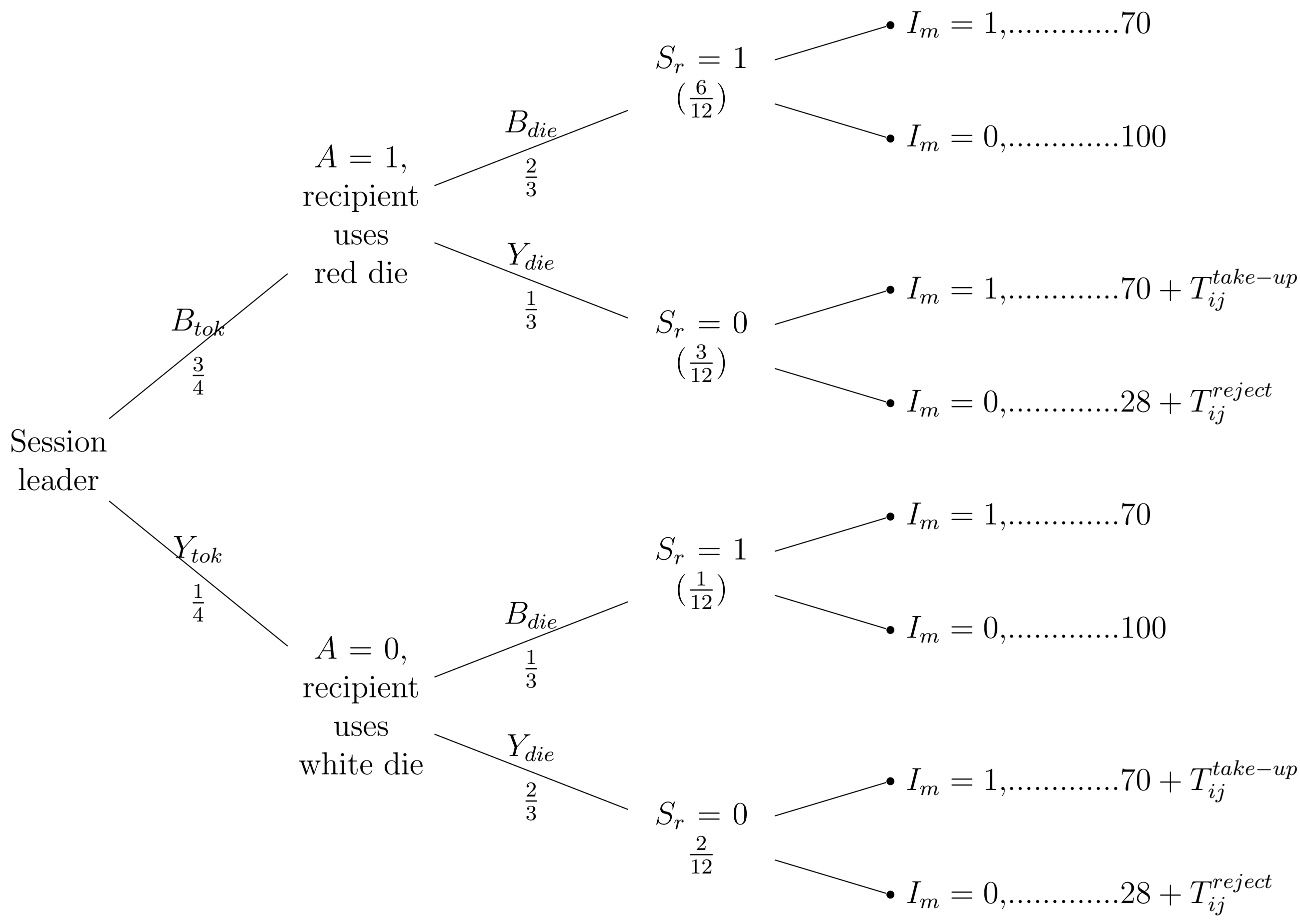

Figure 2.7: Probability distribution and payoff structure in Trt2; B_{tok} represents drawing a blue token, Y_{tok} a yellow token. B_{die} represents a blue outcome from the die roll, Y_{die} a yellow outcome. $A = 0$ represents an aggregate shock, $S_r = 0$ represents an idiosyncratic loss and I_m represents indemnity insurance, if the recipient purchased the insurance $I_m = 1$, otherwise $I_m = 0$

Trt2 subjects' level of understanding

As previously mentioned, we tested both donors and recipients on whether thet understood the game. The composite level of understanding in Trt2 is 98.47% and the percentage level of understanding in each of the 8 questions testing for understanding in this treatment are given below. The top row refers to the questions, numbered as they appear in the questionnaires (see appendix). Below each question is the percentage of farmers who gave the correct response.

$ET2.1$	$ET2.2\_1$	$ET2.2\_2$	$ET2.3\_1$	$ET2.3\_2$	$ET2.4\_1$	$ET2.4\_2$
99.74	99.74	97.19	99.49	97.70	97.70	97.70

2.2.4 The index insurance risk-sharing treatment (Trt3)

The index insurance treatment is identical to Trt2 except that in step 1, prior to the realization of an aggregate shock, a recipient is offered actuarially fair index insurance with premium m_x and payout $P_x = 72$.

As shown in Eq. 2.1, for insurance to be actuarially fair, the premium rate must be equal to the probability of a loss. Index insurance does not provide full indemnity or cover of an actual loss, but pays out on condition of an aggregate shock. Therefore, the premium rate is equal to the probability of an aggregate shock, which from the probability distribution of a Trt3 recipient given in Fig. 2.8 is $\frac{1}{12} + \frac{2}{12} = \frac{3}{12}$. Therefore:

$$m_x = \frac{1}{4} * P_x = 18, \text{ where } P_x = 72, \text{ if } A = 0$$

As in Trt2, it was explained to the participants that in case they chose to take up insurance the premium would be deducted from their income in the game while if they chose not to take up the insurance the premium would not be deducted. Recipients made insurance decisions privately.

The probability distributions and possible outcomes for recipients in Trt3 are presented in Fig. 2.8. Similarly to Trt1 and Trt2, a donor's transfer, T_{ij} is contingent on a recipient experiencing an idiosyncratic loss. In addition, similarly to Trt2, a donor made transfer decisions $T^{take-up}$ and T^{reject} conditional on the recipient's insurance decision. These conditional transfers are added to the end of the branches along which an idiosyncratic loss occurs. The actual transfers were made if a recipient experienced an idiosyncratic loss. If a recipient had taken up the insurance in step 1, then an amount $T_{ij}^{take-up}$ was transferred otherwise an amount T_{ij}^{reject} was transferred.

Trt3 subjects' level of understanding

Both recipients and donor were tested for their level of understanding of the game. The composite level of understanding in Trt3 is 96.67% and the percentage level of understanding in each of the 10 questions testing for understanding in this treatment are given below. The top row refers to the questions, numbered as they appear in the questionnaires (see appendix). Below each question is the percentage of farmers who gave the correct response.

$ET3.1$	$ET3.2\_1$	$ET3.2\_2$	$ET3.3\_1$	$ET3.3\_2$	$ET3.4\_1$	$ET3.4\_2$	$ET3.5\_1$	$ET3.5\_2$	$ET3.6$
100	98.94	94.18	99.21	93.12	99.21	94.18	94.18	94.18	99.47

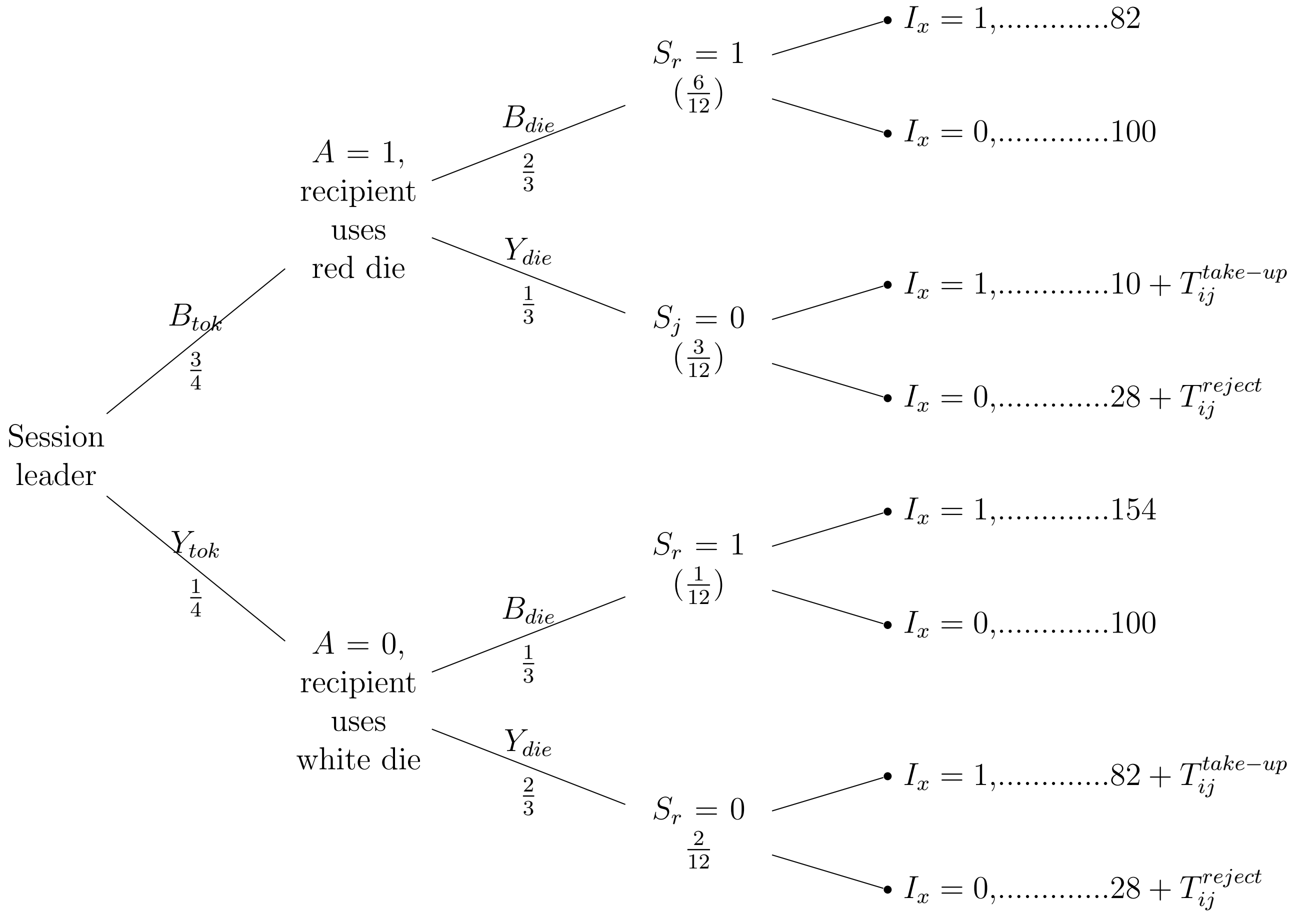

Figure 2.8: Probability distribution and Trt3 payoff structure; B_{tok} represents drawing a blue token, Y_{tok} a yellow token. B_{die} represents a blue outcome from the die roll, Y_{die} a yellow outcome. $A = 0$ represents an aggregate shock, $S_r = 0$ represents an idiosyncratic loss and I_x represents index insurance, if a recipient purchased the insurance $I_x = 1$, otherwise $I_x = 0$.

Chapter 3

Data and exploratory analysis

In this chapter we present the data and discuss factors of internal validity, that is whether we can be sure that the observed changes can be attributed to the experimental treatments and not to other possible causes, like pre-existing differences in the treatment groups [?]. We discuss how we handled missing data and tested for balance in farmer characteristics across the treatment groups. In the final section, we employ plots in exploratory analysis to inform the selection of a suitable statistical analysis and modelling approach.

3.1 Data quality

As discussed in Chapter 2 on sampling and experimental design, randomisation took place at the level of the administrative units, the villages, the iddirs and the farmers with respect to sampling. Additionally, with respect to the experiment, we randomly allocated farmers to treatment groups, pairs and enumerators. We therefore have no reason to believe that there are systematic differences between the treatment groups prior to the effect of the treatment. We verify this using data from the observational study or household survey through Analysis of Variance (ANOVA) to test for the equality of means, of several different farmer or household characteristics that are pertinent to social and economic decision making, across the four treatment groups. This confirmation is necessary for internal validity.

We test the hypotheses $H_o : \mu_{Trt1} = \mu_{Trt2} = \mu_{Trt3}$ against H_a : *Not all treatment means are equal* where the rejection rule is to Reject H_o if the *p-value* < 0.05

Characteristics		Trt1	Trt2	Trt3	$p - value$
Nr of donor farmers		192	194	186	
Age		41.01	41.20	41.30	0.976
Gender	Man	102	110	107	0.955
	Woman	86	82	78	
Marital	Divorced	20	18	23	0.993
	Married	144	151	143	
	Separated	7	12	3	
	Single	4	3	5	
	Widowed	13	7	11	
Household head	No	9	12	11	0.842
	Yes	179	180	174	
Nr of Adults		2.35	1.66	1.88	0.000178***
Nr of Children		3.30	3.16	2.99	0.241
Literate	Yes	106	99	92	0.238
	No	82	93	93	
Education	Adult education	5	4	2	0.018*
	High school complete	3	2	2	
	High school incomplete	3	3	2	
	Junior Complete	4	2	1	
	No education	111	97	88	
	Primary complete	8	19	25	
	Primary incomplete	43	42	56	
	Religious education	11	19	7	
Occupation	Farmer	176	182	165	0.0424*
	Labourer; Farmer	12	3	8	
	Trader; Farmer	0	5	9	
Wealth	Tsem	2.26	2.29	2.39	0.686
	Camels	0.03	0.00	0.02	0.477
	Bulls	0.37	0.35	0.48	0.11
	Oxen	0.89	1.0	0.99	0.406
	Mules	0.01	0.01	0.01	0.999
	Donkeys	0.66	0.69	0.71	0.809
	Horses	0.03	0.01	0.01	0.409
	Sheep	0.87	0.78	0.81	0.921
	Goats	0.88	0.77	0.75	0.895
	Chickens	3.48	3.68	3.96	0.567
	Beehives	0.43	0.72	0.40	0.745
Sharing norms	idd˙yrs	8.41	7.69	8.76	0.671
	idd˙shocK	80.41	58.311	77.55	0.759
	idd˙monthly	7.14	6.59	6.16	0.475
Risk perceptions	25% idio risk exposure	20.64	21.56	21.62	0.728
	50% idio risk exposure	26.70	25.52	26.00	0.712
	25% aggr risk exposure	24.73	27.29	25.89	0.251
	50% aggr risk exposure	34.73	35.99	34.77	0.653

Table 3.1: Summary statistics of the sample characteristics and p-values for ANOVA tests of equality of means across the treatment groups

The results of the ANOVA tests are presented in table 3.1, most farmer characteristics are balanced across the treatment groups. It is not unreasonable that there are a few variables that are unbalanced across treatments, this can be attributed to random chance as by definition, we would expected 5% of the characteristics to turn out to be different. However, in light of the noted highly significant differences in the number of adults per household between treatments, we will include the variable Nr.Adults in the model building process and retain it if found to be significant. The treatments also seem unbalanced with respect to Education and Occupation, hence, these will be used to check for robustness.

3.2 Missing data

Missing data may be a problem which if dealt with inappropriately can lead to bias or compromise statistical efficiency by increasing the standard errors and reducing the power of a test. There are three types of possible missingness, namely, Missingness Completely At Random (MCAR), Missingness At Random (MAR) and Missingness Not At Random (MNAR). Data is MCAR if it is completely independent of the experiment, for instance, when measuring equipment fails in an experiment [**?**]. In our case several questionnaires were reported missing at the data capturing stage. This is MCAR.

Data is MAR if it is independent of the unobserved data. This means the missing data contains no information about the unobserved data points that is not already contained in the observed data points. For instance in a longitudinal study whereby some measure of illness is taken repeatedly but the patient recovers and observation stops. Data is MNAR if it is dependent on the unobserved data. For instance, when a patient fails to attend data collection because they are too ill or have recovered but the researcher does not have this information. MCAR is less problematic, discarding the missing cases, by list-wise deletion does not lead to bias although statistical power may be lost. On the other hand, MAR requires judicious use of the observed values in order to avoid bias. MNAR is much difficult because the missingness pattern/distribution contains information about the unobserved data points, but the researcher may not know which information, thus resolving the problem of MNAR requires additional assumptions, which further increases the uncertainty of conclusions drawn from the analysis. [?],

Due to randomly missing data (MCAR), our sample size reduced to 1527 farmers. However, mixed-effects models are able to handle unbalanced data and produce efficient estimates [??]. Hence we proceeded to model fitting without employing any other techniques.

3.3 The data: measures of central tendency

In all treatments, we asked how much a donor farmer would transfer to a recipient farmer if the recipient suffered a loss. This was recorded as Tij_{GEN}. This is an unconditional transfer, which reflects a transfer made when insurance is available, hence for analytical purposes, we rename it T_{ij}^{avail}. Technically speaking, in Trt1, the more appropriate by-name is $T_{ij}^{unavail}$. In Trt2 and Trt3 we further asked how much a donor would transfer on condition that the recipient took up the insurance and on condition that the recipient did not take up the insurance. For our analysis these are $T_{ij}^{take-up}$ and T_{ij}^{reject} respectively. Note that when the donor makes a transfer decision, he or she does not know if the recipient has taken up or rejected the insurance. It is also unknown to the donor whether the recipient has actually experienced an idiosyncratic loss. Table 4.1 summarises the analytical response variables. Although it is possible to conduct a within-subject analysis based on Trt4 observations, this thesis restricts its analysis to a between-subject comparison, hence makes use of only Trt1, Trt2 and Trt3 observations.

In each treatment, we have two possible pair-types, namely, PTd and PTs. We also have two possible weather outcomes based on the realization of an aggregate shock. B (blue token drawn) when there is no aggregate shock and Y (yellow token drawn) when there is an aggregate shock. Hence, from the 4 experimental conditions we can create 4 subsamples, namely, $BPTs$, $BPTd$, $YPTs$ and $YPTd$. We present the transfers per treatment for the subsamples under the three possibilities of insurance availability, take-up and rejection in Table 3.2.

			$T_{ij}^{unavail}$	T_{ij}^{avail}		$T_{ij}^{take-up}$		T_{ij}^{reject}	
		obs	Trt1	Trt2	Trt3	Trt2	Trt3	Trt2	Trt3
B	PTd	261	17.31	19.54	13.61	5.18	14.41	16.71	9.13
			(8.71)	(10.16)	(9.19)	(6.48)	(10.31)	(12.2)	(8.48)
	PTs	282	17.01	19.10	16.05	5.72	16.52	15.43	11.76
			(8.28)	(10.18)	(8.54)	(7.07)	(11.30)	(10.73)	(9.95)
Y	PTd	80	12.11	20.36	13.15	7.00	3.85	16.20	12.08
			(5.85)	(9.40)	(5.24)	(8.66)	(4.16)	(9.54)	(7.43)
	PTs	78	18.13	21.72	15.15	8.12	6.15	19.04	15.31
			(8.79)	(9.82)	(7.49)	(8.57)	(6.82)	(10.71)	(9.30)

Table 3.2: The mean (and standard deviation in parentheses) of the observed transfers in ETB under insurance unavailability in Trt1, availability, take-up and rejection in Trt2 and Trt3 for the 4 subsamples. The column obs is the number of observations in that subsample.

Farmers from different iddirs (PTd) seem to transfer more when there is no aggregate shock than when there is an aggregate shock, particularly in Trt1 and Trt3. For example, in Trt1 subsample B with no aggregate shock, hence sample $BPTd$ transfer an average of 17.31 ETB. On the other hand, when there is an aggregate shock, farmers YPTd transfer an average of 12.11 ETB. In contrast, farmers from the same iddir transfer more when there is an aggregate shock than when there is no aggregate shock, particularly in Trt1 and Trt2. Farmers in subsample BPTs in Trt2 transfer an average of 19.10 ETB while farmers in YPTs in Trt2 transfer an average of 21.72.

For indemnity insurance, comparison of the means of the conditional transfers shows that a recipient's take-up of indemnity insurance decreases a donor's transfer, that is, $T_{ij}^{uptake} < T_{ij}^{avail}$. Similarly, a recipient's rejection of indemnity insurance decreases a donor's transfer. However, the reduction under rejection is less substantial than the reduction under take-up; that is $T_{ij}^{take-up} < T_{ij}^{rejection} < T_{ij}^{avail}$.

For index insurance, comparison of the means of the conditional transfers shows that when there is no aggregate shock, a recipient's take-up of index insurance increases a donor's transfer only marginally, even for PTs. However, a recipient's rejection of index insurance decreases transfers more substantially. Hence, $T_{ij}^{rejection} < T_{ij}^{avail} < T_{ij}^{take-up}$. On the other hand, when there is an aggregate shock a recipient's take-up of index insurance reduces a donor's transfer substantially, whereas rejection results in a marginal reduction, that is $T_{ij}^{take-up} < T_{ij}^{rejection} < T_{ij}^{avail}$.

3.4 The data: distribution of the data

We explore the distribution of the transfers using histograms. Transfers under take-up and rejection exhibit multimodality regardless of aggregate shock. Fig. 3.1 shows transfers when there is no aggregate shock while Fig. 3.2 shows transfers when there is an aggregate shock.

When indemnity insurance is available, transfers peak at 0, 10 and 20 ETB, reaching the highest frequency at 30 ETB. However, under take-up the most frequently transferred amount is 0 ETB, while under rejection, peaks at 0, 10 and 30 ETB are somewhat equal. When there is an aggregate shock, transfers exhibit a similar pattern.

For index insurance, the most frequent transfer is 10 ETB, this is consistent when index insurance is available, taken up or rejected. When there is an aggregate shock, transfers exhibit a similar pattern.

3.5 The data: multilevel structure with crossed factors

In our study, farmers are sampled from iddirs, iddirs from villages and villages from administrative units, hence we have four levels of hierarchy in the sampling design. Levels are numbered according to the most detailed; hence the farmers are level-one units whilst the iddirs are level-two units and so on. We limit our analysis to these two levels because the objective of the study is to understand farmers' transfer behaviour as a result of their

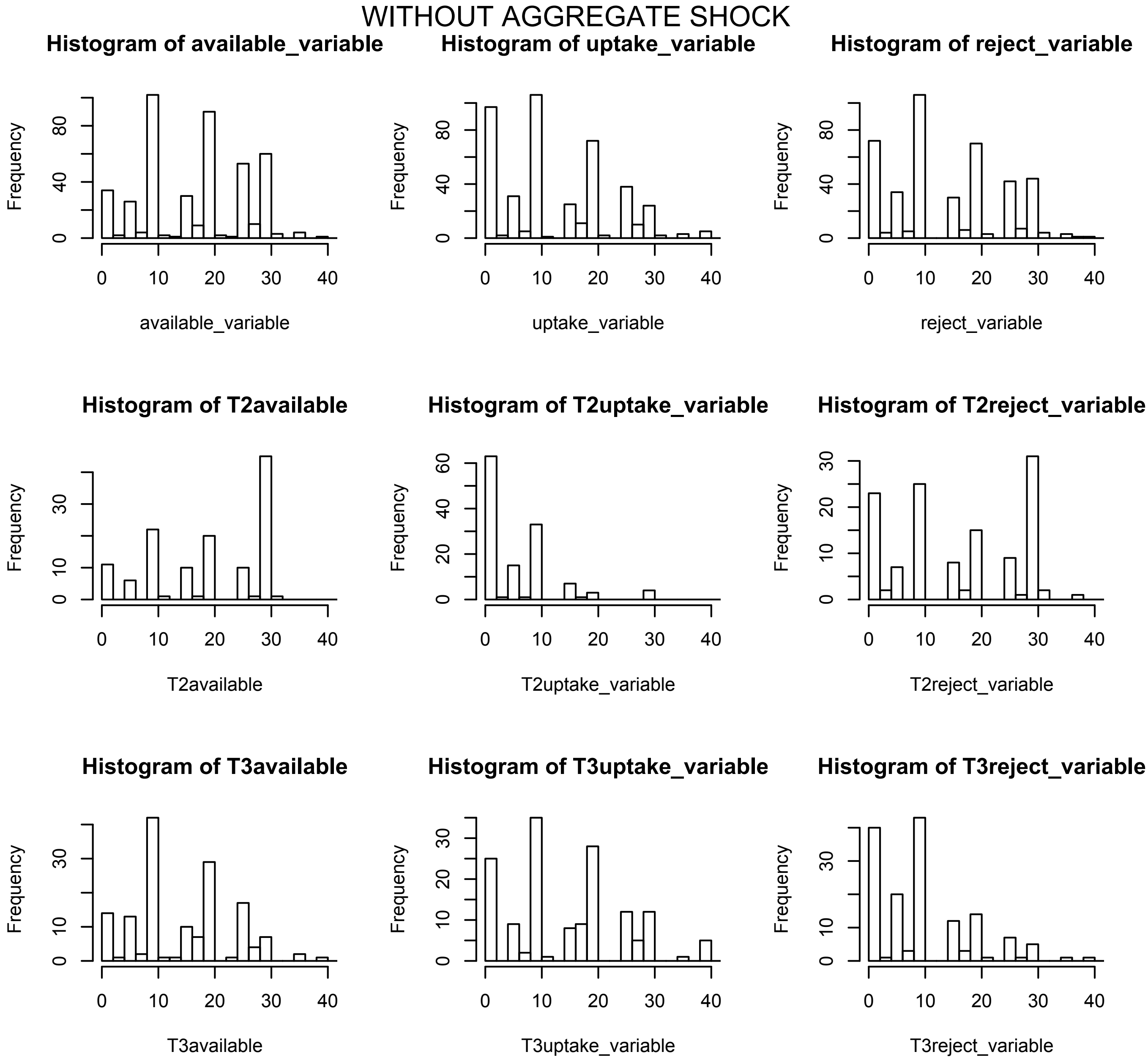

Figure 3.1: The distribution of transfers when there is no aggregate shock, under insurance availability, take-up and rejection. Transfers in both treatments exhibit multimodality.

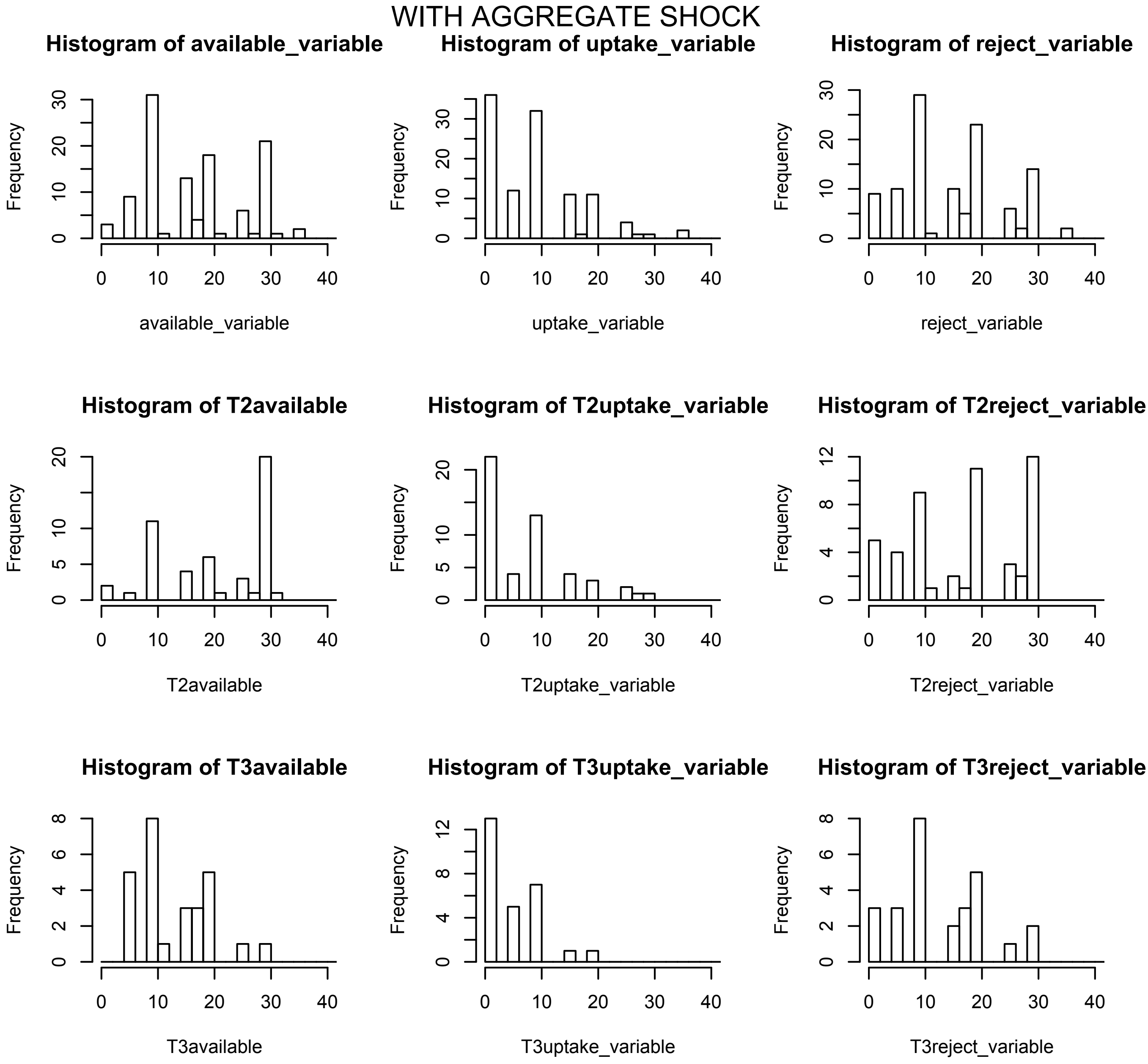

Figure 3.2: The distribution of transfers when there is an aggregate shock, under insurance availability, take-up and rejection. Transfers in both treatments exhibit multimodality.

sharing norms, which are exhibited at iddir level. Furthermore, based on statistical theory, we assume that with only 3 administrative units and 2 to 3 villages, we do not have enough levels of these categorical variables to model them as random effects. That is, the level of variability between only 3 administrative units and only 2 or 3 villages does not warrant incorporating random effects in the model. In addition, at level two, we have an enumerator categorical variable with 10 levels as the experiment was conducted with 10 enumerators.

The administrative units were not randomly sampled, thus the external validity of the study is limited in the sense that a similar study using farmers sampled from different areas of Tigray province may not produce the same results. We are thus unable to generalize beyond the administrative units that we sampled from. However, we expect that the mechanistic findings would be similar beyond these administrative units, across rural Tigray.

We define factors as classifications which identify each source of the datum in our data. There are several classes within each factor and the points at which each class of a factor intersects with each class of another factor are called cells [**?**]. A dataset with equal observations or datum in each cell is referred to as *balanced*. However, as is the case with most experiments, our data consists of empty cells and is therefore unbalanced. The treatment factor has 3 levels (*Trt1, Trt2 and Trt3*), the iddir factor has 16 levels (*idd1, idd2, . . . , idd16*), the enumerator factor has 10 levels (*Abr, Ashu, Beky, Gech, Gir, Haf, Kib, Mul B, Mul K, Samy*). In each iddir the set of 6 farmers in Trt1 is different from the set of 6 farmers in Trt2 and Trt3. Furthermore, as illustrated in Fig 2.3, within each iddir, in each pair-type, we have 6 Trt1 pairs but because the 6 Trt1 farmer pairs in *idd1* are different or independent of the 6 Trt1 farmer pairs in *idd2* and so on, treatment farmer pairs are *nested* within the iddir factor. Concurrently, the same set of 10 enumerators is used in all iddirs, hence as shown in 2.5, the enumerator and iddir factors are *crossed*. We have at least one observation for each combination of a level of enumerator and iddir therefore the factors are *completely* crossed [**?**].

This structure may affect our measurement of interest, the transfer, because some iddirs may have higher transfers than others. Iddirs may differ in their familiarity or experience of either index or indemnity insurance, creating systematic differences in their levels of transfers. Although the enumerators received the same training to standardize the experiment, their different personalities may have translated into differences in explaining the risk sharing game. All these factors may affect the level of transfers we measure, thus we control for this additional variation by estimating enumerator and iddir random effects.

3.6 Exploratory data analysis

So far we summarised the data by its measures of central tendency and viewed its general multimodal distribution by plotting histograms. In this section we conduct analysis using plots because they provide visual summaries which offer a more complete description of the data than simple statistical measures. We use exploratory data analysis as an inductive tool, seeking patterns in the data to guide the selection of an appropriate modeling procedure [**?**].

Plots comparing transfers when insurance is not available to transfers when insurance is available, taken up and rejected are given in Fig. 3.3. Generally, when insurance is not available, transfers are positively skewed, have the widest spread and highest average. Farmers are better off when indemnity insurance is available, transfers have the highest mean and smallest spread. In contrast, the widest spread and lowest mean occurs when index insurance is available, which suggests that farmers may be worse off under this condition. Next we compare transfers taking into account pair-type and aggregate shock.

Fig 3.4 compares the transfers by pair-type and aggregate shock. Transfers for pairs from different iddirs are on average similar to transfers for pairs from same iddirs. However, under insurance rejection, donors transfer slightly less to recipients from different iddirs. When there is no aggregate shock, transfers have more spread and are more positively skewed than when there is an aggregate shock. As expected, insurance rejection decreases transfers independently of the aggregate shock, whereas the realization of an aggregate shock reduces transfers more under insurance take-up. There also seem to be outliers under insurance take-up and rejection. Next, we explore the transfers with respect to iddirs.

A plot of the transfers with respect to iddirs in Fig 3.5 shows that a farmer's transfer decision is affected by the iddir that they are a member of. Iddir 2 has particularly low mean transfers across all three analytical possibilities. We are not interested in knowing specific effects of individual iddirs, hence we will model iddir as a random effect. This is discussed in subsequent sections. We are similarly uninterested in the specific effects of the individual enumerators, hence we will also have random enumerator effects. On average *Gir* recorded lower transfers, the effects of the other enumerators are displayed in Fig 3.6.

From Fig 3.4, there is evidence that the aggregate shock also has an effect on a donor's transfer decision. From Fig 3.5 and Fig 3.6, we have seen that iddir and enumerator effects matter, as does Pair-Type, particularly under insurance rejection. Hence we stratify the data into two samples; B, which is the sample of farmers who did not experienced an aggregate shock and sample Y, which consists of farmers who experienced an aggregate shock. Thus we present consolidated plots showing these effects on the mean of the transfers under three possibilities of insurance availability, take-up and rejection, contrasting sample B to sample Y.

Under insurance availability, which is shown in Fig 3.7, with no aggregate shock, pair-type is nearly insignificant, whilst with an aggregate shock, different iddir pair-types transfer much less. Treatment effects also seem to differ, Trt1 transfers are more than Trt3 transfers with no aggregate shock but less with an aggregate shock. Additionally, iddir effects vary due to aggregate shock. For example, iddir 8 has the highest transfer with no aggregate shock but the lowest with an aggregate shock. Finally, the enumerators also seem to be affected by the aggregate shock, $Mul\,B$ has the lowest transfers with no aggregate shock but with an aggregate shock he records the third highest transfers.

Under insurance take-up, which is shown in Fig 3.8 the aggregate shock realized interacts with pair-type in an opposite way from the interaction under insurance availability, we see negligible differences with an aggregate shock and larger differences with no aggregate shock. We also see differences in transfers between Trt2 and Trt3 with no aggregate shock range between $(4.5 - 7.5)$ whereas with an aggregate shock, the differences range between

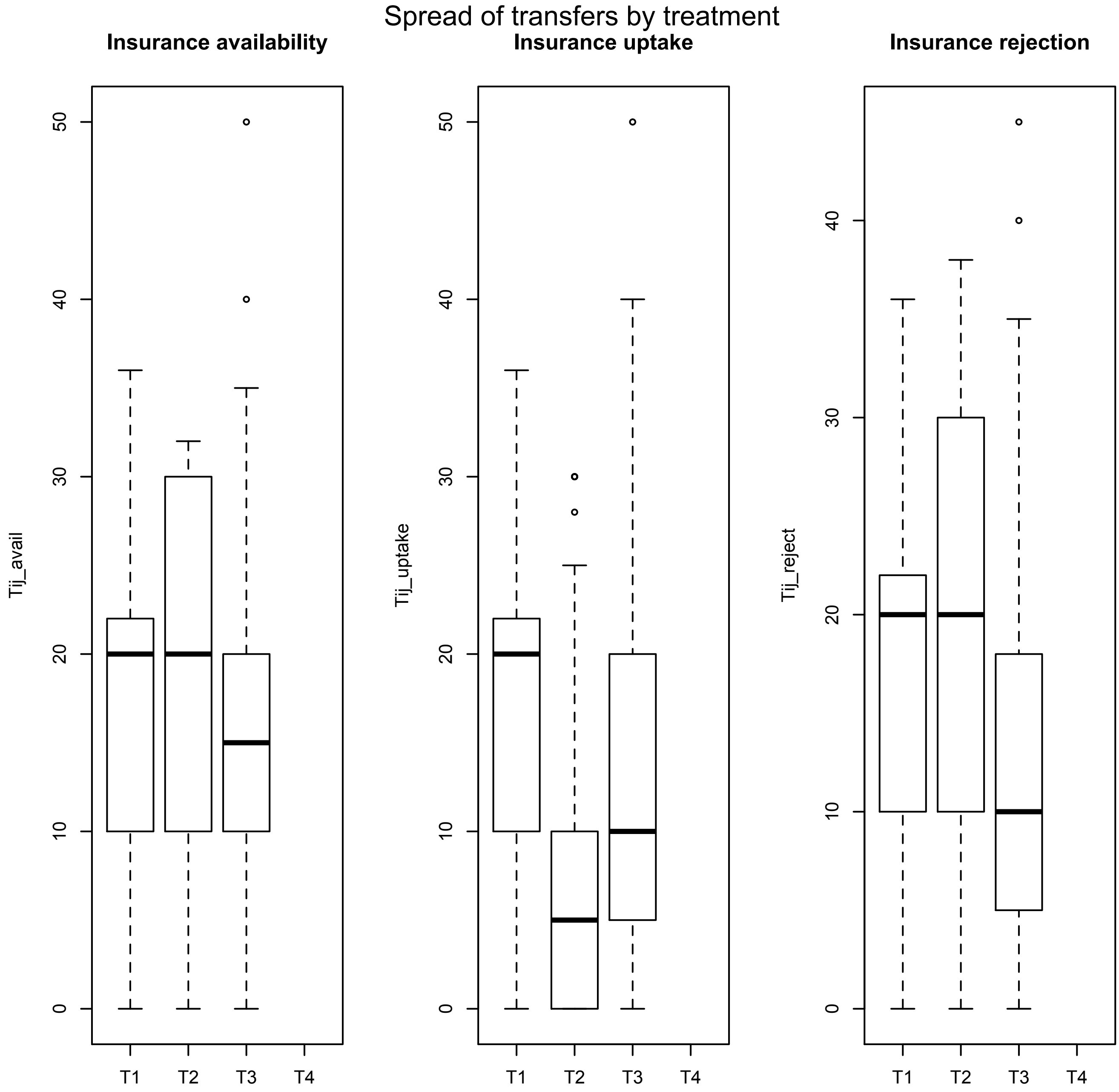

Figure 3.3: The spread of transfers by treatment

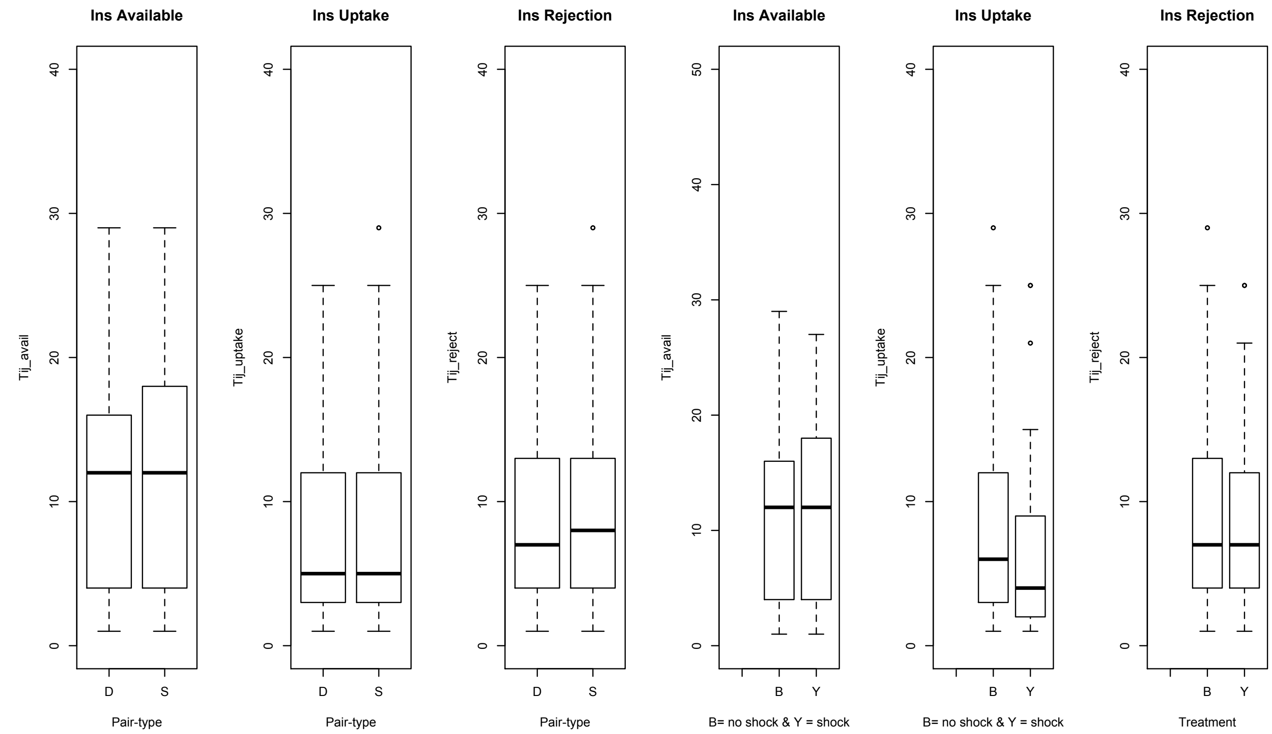

Figure 3.4: Transfers by pair-type and aggregate shock

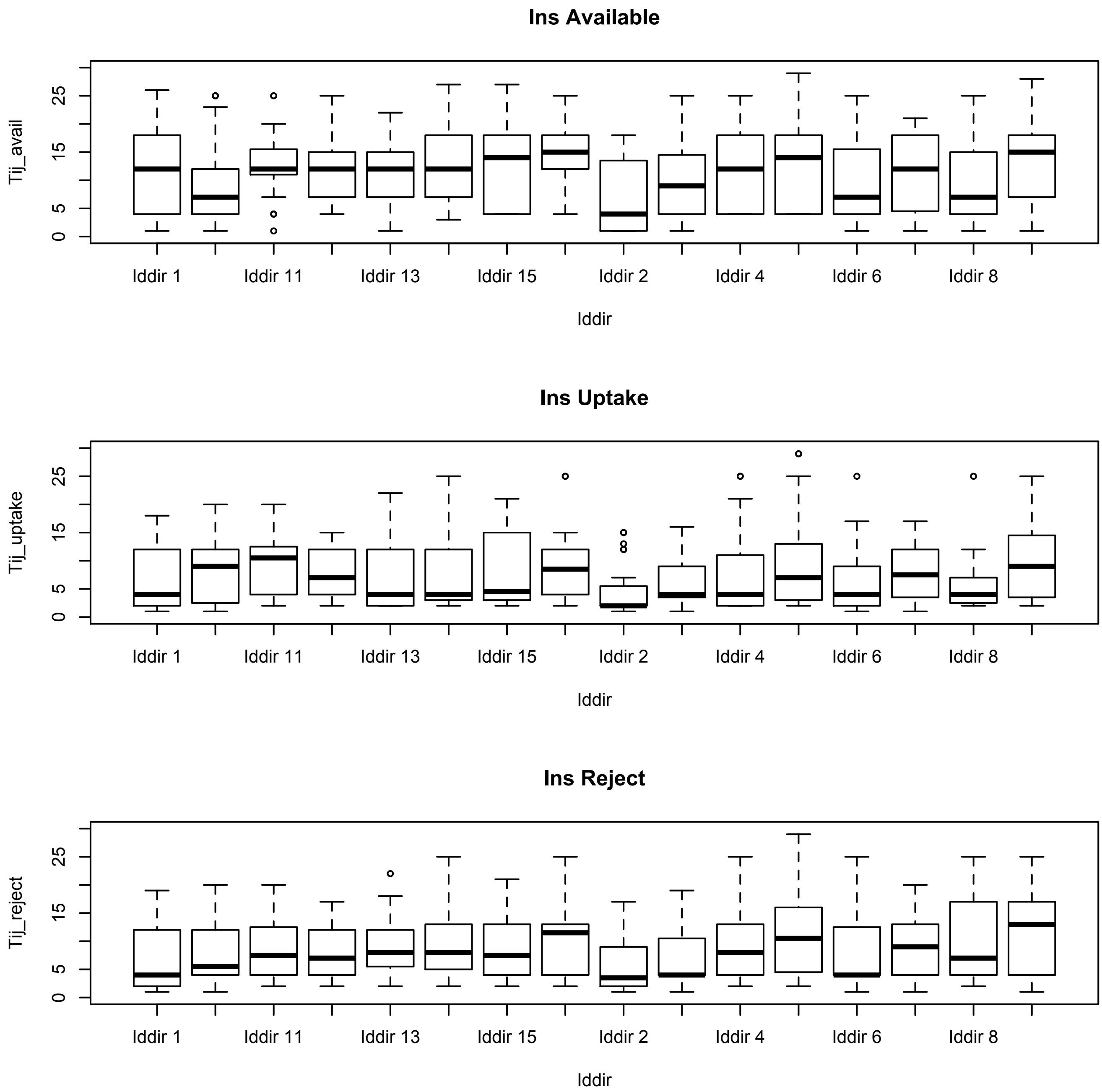

Figure 3.5: Transfer under insurance availability, take-up and rejection by iddir

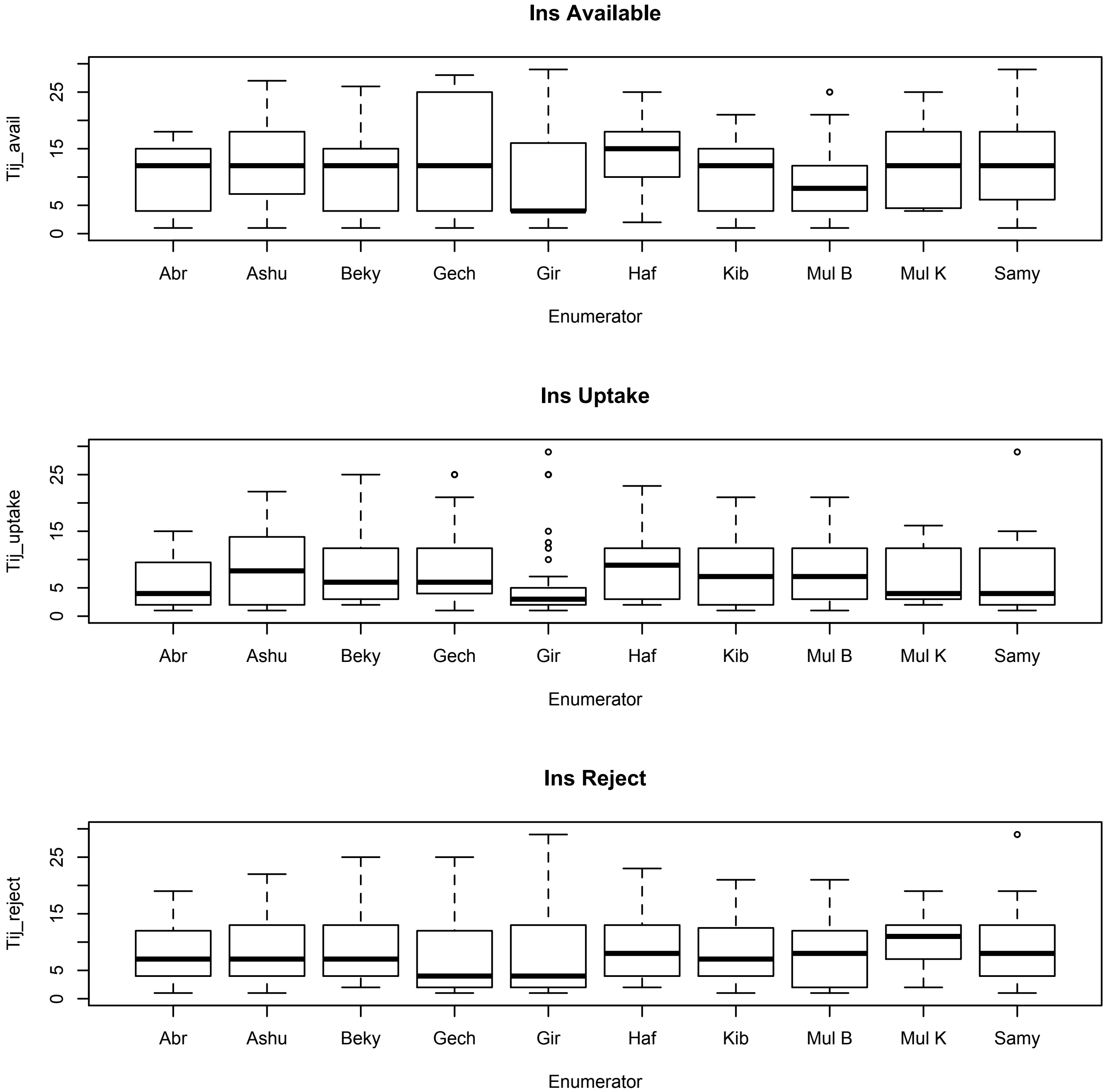

Figure 3.6: Transfer under insurance availability, take-up and rejection by enumerator

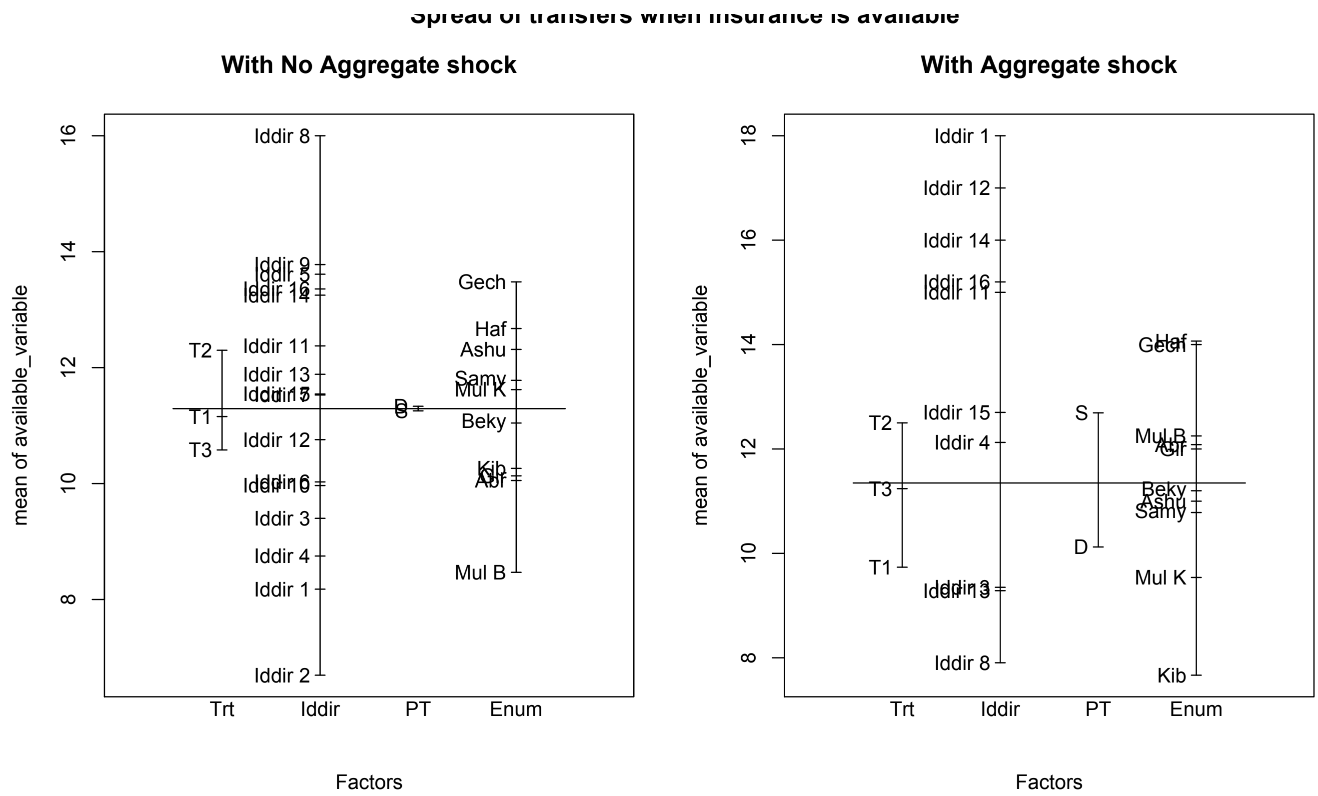

Figure 3.7: Spread of transfers when insurance is available

(4.5 – 5.5). Similar to the case under insurance availability, iddir and enumerator effects are different with an aggregate shock and without an aggregate shock, for example, enumerator *Ashu* has the highest transfers with no aggregate shock, but the lowest with an aggregate shock.

We also see evidence of differences in interactions of aggregate shock with treatment, pair-type, iddir and enumerator under insurance rejection, which is shown in Fig 3.9.

Analysis in this section shows that in order to explain how donors' transfers change in response to recipients' decisions to take up or not to take up insurance it would be useful

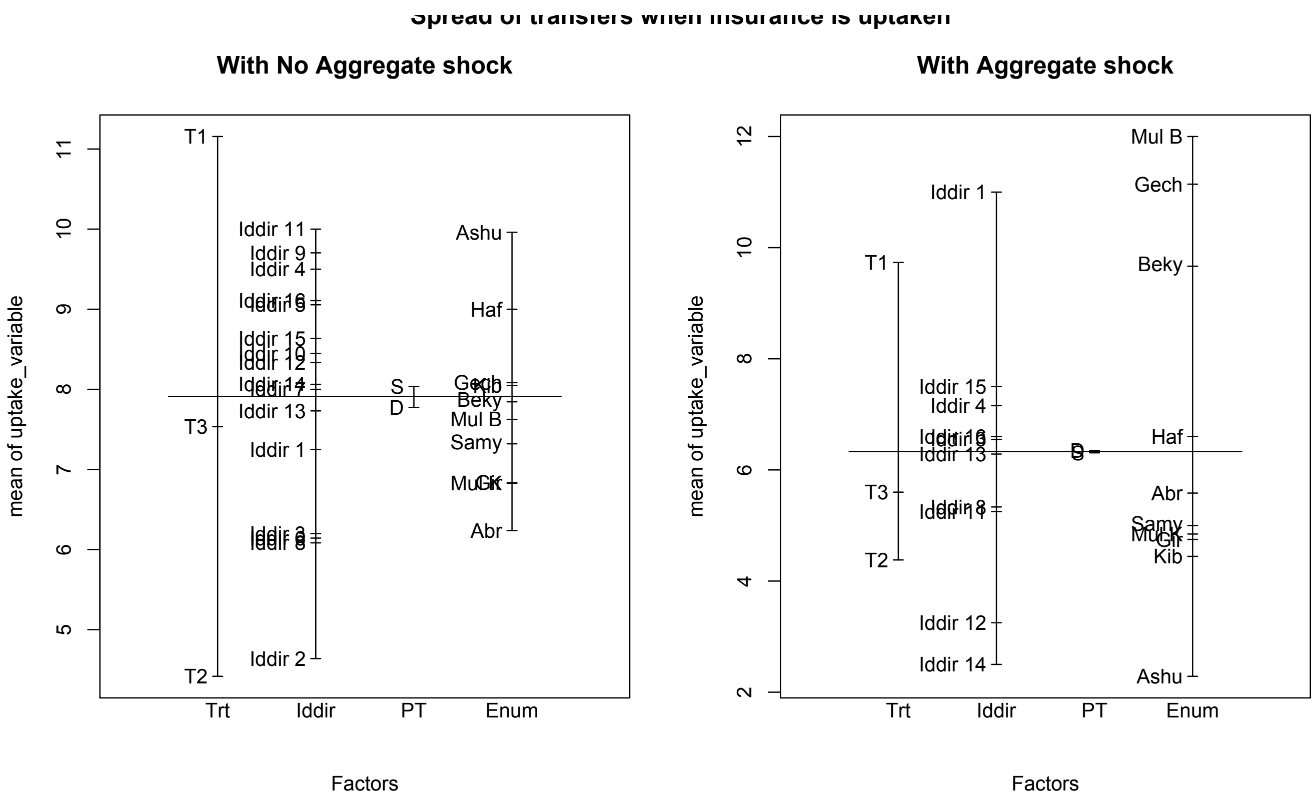

Figure 3.8: Spread of transfers when insurance is taken up

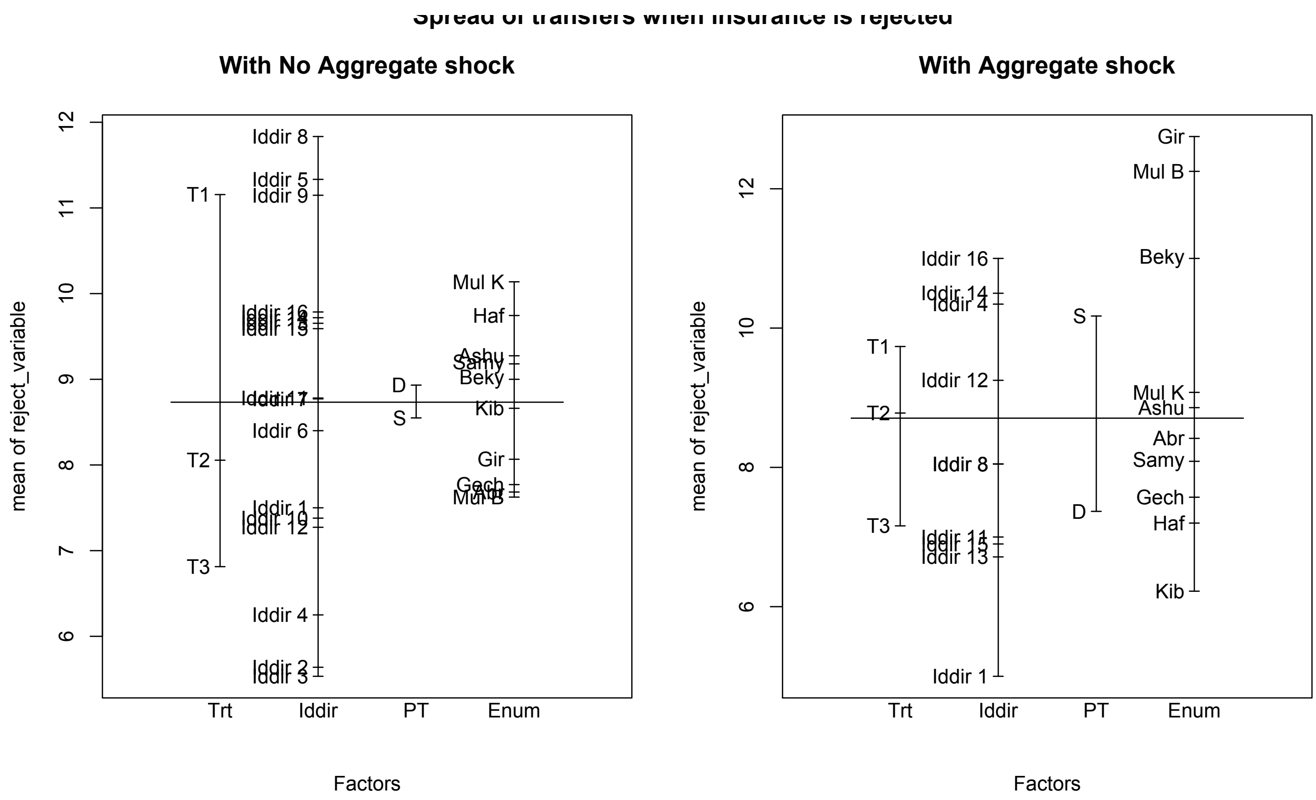

Figure 3.9: Spread of transfers when insurance is rejected

to stratify the sample by aggregate shock as well to account for iddir and enumerator effects. We now turn to this statistical analysis.

Chapter 4

Statistical Analysis

In this chapter we discuss the Generalized Linear Model because the data cannot be assumed normal. Furthermore, its hierarchical nature results in correlation between errors, hence we also discuss the Generalized Linear Mixed Model. Our response variable(s) are discrete interval data and our main explanatory variable is an unordered categorical variable, namely a treatment factor with 3 levels, Trt1 (risk-sharing with no insurance), Trt2 (risk-sharing with indemnity insurance) and Trt3 (risk-sharing with index insurance). We use different response variables to analyse the treatment effect under the three possibilities of insurance availability, take-up and rejection. Table 4.1 presents the analytical variables, whose means are given in Table 3.2. We also have secondary explanatory variables, which are a mixture of continuous, categorical and discrete variables. A full list of the variables is provided in the appendix, of which the ones of interest are pair-type, enumerator and iddir.

Recorded variable	Analytical variable	Trt1	Trt2	Trt3
Tij_GEN	$T_{ij}^{unavail}$	✓		
Tij_GEN	T_{ij}^{avail}		✓	✓
Trt1ij_S		✓		
Trt1ij_D		✓		
Tij_No_Ins	$T_{ij}^{unavail}$	✓		
Tij_C_Y	$T_{ij}^{take-up}$		✓	
Tij_C_No	T_{ij}^{reject}		✓	
Tij_R_Y	$T_{ij}^{take-up}$			✓
Tij_R_No	T_{ij}^{reject}			✓
Like_C_Ins			✓	
Like_R_Ins				✓

Table 4.1: Table of analytical response variables

4.1 Generalized Linear Models

In this section we discuss the peculiarities of our data that necessitate the use of generalized linear mixed models rather than the traditional linear model. Thereafter we describe a generalized Linear Model (GLM), its measures of goodness of fit and diagnostics.

There are four main reasons why a linear model may not be appropriate for our data. Firstly, a linear model may predict negative transfers, which would mean that the model is inaccurate. Secondly, errors may not be normally distributed, due to the grouped nature of the data as a result of multi-sampling. Thirdly, variance may be non-constant heteroscedastic) n. The exploratory analysis shows that the variance-mean ratio for most of the analytical variables is less than one, most of the means are greater than the variance except for $T_{ij}^{take-up}$ in Trt2 (see Table 3.2). Finally, transformations that are easy to handle with a linear model may be difficult with values of zero [?].

4.1.1 Description of a Generalized Linear Model

A Generalized Linear Model (GLM) is defined by two components; an exponential family distribution that the response is assumed to follow and a link function that describes the relationship between a linear combination of the predictor variables and the mean of the response. The linear combination of predictors is known as the *linear predictor*, often denoted η.

The exponential families

The exponential family of distributions, which the response variable Y in a GLM behaves like, is of the general form

$$f(Y = y|\theta, \phi) = \exp\left(\frac{y\theta - b(\theta)}{a(\phi)} + c(y, \phi)\right), \tag{4.1}$$

with mean $E(Y) = \mu = b'(\theta)$, and variance $var(Y) = b''(\theta)a(\phi)$, where θ is called the *canonical parameter*, which represents the location while ϕ is called the *dispersion parameter*, which represents the scale. Different specifications of the functions a, b and c results in specific exponential family distributions.

The Normal or Gaussian distribution arises when $\theta = \mu$, $\phi = \sigma^2$, $a(\phi) = \phi$, $b(\theta) = \theta^2/2$ and $c(y, \phi) = -\frac{1}{2}(\frac{y^2}{\phi} + \log(2\pi\phi))$, to give:

$$\begin{aligned} f(y|\theta, \phi) &= \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(y-\mu)^2}{2\sigma^2}\right) \\ &= \exp\left(\frac{y\mu - (\mu^2/2)}{\sigma^2} - \frac{1}{2}(\frac{y^2}{\sigma^2} + \log(2\pi\sigma^2))\right). \end{aligned} \tag{4.2}$$

Other exponential distributions are the binomial, which is commonly used for binary count data and the Poisson, commonly used for count data. Less common are the gamma and the inverse Gaussian distributions. In addition, the negative binomial and the Weibull distributions are sufficiently close that with some modifications, they can be used to fit GLMs although they are not part of the exponential families of distributions [**?**].

The linear predictor

The effect of the predictors on the response is expressed through η, a *linear predictor*:

$$\eta = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_p x_p = \mathbf{X}\boldsymbol{\beta},$$

which in a Generalized Linear Mixed Model (GLMM) includes both fixed and random effects such that

$$\eta = \beta_0 + \beta_1 x_1 + \cdots + \beta_p x_p + b_1 z_1 + \cdots + b_q z_q = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{b}.$$

$\mathbf{X}$ and $\mathbf{Z}$ are model matrices that consist of covariates, while $\boldsymbol{\beta}$ is a vector of fixed parameters and $\mathbf{b}$ a vector of random effects. We return to this in the section on GLMMs.

The link function

The link function g connects the mean of the response to the linear predictor:

$$\eta = g(\mu).$$

Although g can be any monotone continuous differentiable function, there are convenient functions for GLMs which are commonly used. In a Gaussian linear model, the identity function is used, which yields $\eta = \mu$. For a Poisson GLM the standard choice is a log function, which ensures that the model does not predict a negative count, even though η may. So $\eta = \log \mu$, therefore $\mu = e^{\eta}$ ensures $\mu > 0$. The log link means that additive effects of x lead to multiplicative effects on μ.

If $\eta = g(\mu) = \theta$, then g is a *canonical link* because it results in the canonical parameter of the exponential distribution. Because $E(Y) = \mu = b'(\theta)$, a canonical link gives $g(b'(\theta)) = \theta$ and if a canonical link is used, then $E(Y)$ is *sufficient* for $\boldsymbol{\beta}$ [?].

4.1.2 Parameter estimation

GLMs are fitted by solving the set of likelihood equations for estimates of the parameters $\boldsymbol{\beta}$. For a Gaussian, an exact solution for the maximum likelihood estimate $\hat{\boldsymbol{\beta}}$ is possible, but for other GLMs, the estimation procedure requires numerical optimization as the likelihood equations are usually nonlinear in the coefficients of the linear predictor. The solutions are then found iteratively. McCullagh and Nelder (1989) applied the Newton-Raphson method with Fisher scoring to show that optimization is equivalent to iteratively reweighted least squares (IRWLS) [?]. While the Newton-Raphson method uses the observed information matrix, the Fisher scoring uses the expected information matrix. However, for canonical link models, these are the same. Therefore it is advisable to use canonical link functions rather than specifying a different link function when fitting a GLM [?].

4.1.3 Goodness of fit

Statistical models are approximations, they therefore are unlikely to fit the data perfectly. Notwithstanding, in model fitting there are two extremes, a *null* model and a *saturated* model. The null model simply fits a common mean, that it, it has one parameter and represents a situation where the response and covariates are not related at all and the data is entirely explained by random variation. In contrast, a saturated model fits as many parameters as there are data points, such that the model is explained exactly. This represents the data as being entirely systematic. The essence of model fitting is to represent both the systematic component and the random component of the data. Hence between the null and a saturated model, there is a range of possible models that can be fitted to the data. Evaluating how good one model is over other potential models is therefore critical [?].

Because a full model gives the measure of how well any model can fit the data, the difference between the log likelihood of the full model $l\left(y, \phi/y\right)$ and the log likelihood of the potential model $l\left(\hat{\mu}, \phi/y\right)$ can be seen as the size of the discrepancy between the model and the data. That is, twice the difference between the log likelihood of the given model and the log likelihood of a perfect fit or saturated model is known as a **likelihood ratio statistic** $\mathcal{LR}$, which is of the general form:

$$\mathcal{LR} = 2(l\left(y, \phi/y\right) - l\left(\hat{\mu}, \phi/y\right)).$$

For independent observations of an exponential family distribution, when $a_i(\phi) = \phi/w_i$, then the $\mathcal{LR}$ simplifies to a measure called the **scaled deviance**:

$$D(y,\, \hat{\mu})/\phi = \sum\nolimits_i 2\, w_i \left(y_i(\tilde{\theta}_i - \hat{\theta}_i) - b(\tilde{\theta}_i) + b(\hat{\theta}_i) \right) / \phi,$$

where $\tilde{\theta}$ and $\hat{\theta}$ are the estimates under the saturated and the proposed model, respectively. $D(y,\, \hat{\mu})$ without a scale factor is the **deviance**.

Although different families of the GLMs estimate deviance differently, the deviance is interpreted similarly, as an indication of the extent of inadequacy of the model. Hence, minimising the deviance is maximising the likelihood. Therefore, the deviance is understood as a measure of goodness of fit.

An alternative measure of discrepancy or goodness of fit is the Pearson's X^2 statistic :

$$X^2 = \sum\nolimits_i \frac{(y_i - \hat{\mu})^2}{var(\hat{\mu})}.$$

The goodness of fit can be used to compare models. This is based on the fact that under certain conditions, if the proposed model is correct, the scaled deviance and the Pearson's X^2 statistic are asymptotically χ^2 with degrees of freedom equal to the number of identifiable parameters. However, the test is only practical if the dataset is large enough, and the dispersion parameter is known, as in the binomial and Poisson where $\phi = 1$. The test is unreliable for a binary response and unsuitable for the Gaussian because ϕ is unknown [**??**]

4.2 Mixed-effects modeling or multilevel analysis

In this section we discuss the peculiarities of modelling data with a grouped structure or hierarchy, which violates the assumption of independent errors that a traditional fixed-effects linear model assumes. Thereafter we describe the distributions and notation of Linear Mixed-effects Models (LMMs).

Mixed-effects or multilevel analysis is one of the basic techniques used for modelling data with either complex hierarchical structures, or experiments with multistage sampling. The objective is to make inferences about social and behavioural mechanisms. Multilevel models are alternatively named hierarchical, mixed-effects or random coefficients models based on their technical elements [?]. In this thesis, we will refer to them interchangeably as mixed-effects models or multilevel models.

Why use a mixed-effects model?

Our aim is to analyse the effect of insurance availability, take-up and rejection. We do this by comparing transfers under these three possibilities to transfers when formal insurance is not available, with all other predictors held constant. In a classical regression we would use interaction terms. However, our data are clustered. Farmers are nested in iddirs due to the sampling design and enumerators and iddirs are crossed due to the experimental design. In a classical regression this would require inclusion of indicators for clusters at all levels and would result in over-fitting if Maximum Likelihood estimators are used [?]. The remaining alternatives, aggregation (or pooling) and disaggregation, which are described briefly in the following discussion are equally unsatisfactory.

One alternative is to aggregate the transfers for each iddir and perform an OLS regression on the averages incorporating indicator terms for each iddir. However, aggregation essentially discards individual farmer-level information and makes causal inference problematic [**??**]. In addition, such an approach would be less parsimonious than a mixed-effects model and because the iddirs are not treated as a random sample, we would be unable to quantify variation between iddirs, such a model would therefore be inefficient.

Another alternative is to ignore the iddir terms and treat the data as if it had no clustering. This approach is known as disaggregation, it involves treating the level one variables as if they are independently drawn observations. By using the characteristics of the higher-level units as characteristics of the lower-level units within them, it ignores the grouped nature of the data and hence exaggerates the sample size. Instead of 16 independently sampled iddirs with 48 donor farmers randomly sampled from each, it would be akin to 768 independent observations. The problem is that the chances of committing a type I error increases when investigating between-iddir relationships. Conversely when studying within-iddir relationships, disaggregation results in type I error probabilities that are too low [**?**]. Mixed-effects modelling enables calculation of correct standard errors by capturing variation at all levels within the data or similarly partitioning the overall variation into iddir, enumerator and treatment components.

4.2.1 Description of a Linear Mixed-effects Model

Statistical models describe a relationship between a *response* variable and *covariates* measured with the response variable. In linear mixed-effects models, at least one of the covariates is categorical and represents the experimental or observational factor in the dataset [?]. In our study the experimental unit is treatment, which is a categorical variable with three levels while, iddirs and enumerators represent the observational units. These are also categorical.

Parameters associated with levels of a covariate are referred to as *effects*. If the levels of a covariate are fixed or finite we model the covariates using fixed effects parameters. Hence, parameters that are constant or reproducible are called *fixed*effects. If the levels of a factor constitute a random sample from the set of all possible levels of that factor, then we model the covariates using *random* effects. Thus the terms *fixed* and *random* refer to levels of the categorical variable or factor rather than the parameters themselves. Therefore, a mixed-effects model incorporates both fixed and random effects [?].

In a linear mixed-effects model, each random-effects term is defined with respect to a categorical covariate, which is most often a grouping factor whose levels are regarded as a random sample from a population of possible levels [?]. However, categorical covariates with fixed or a finite number of levels, for example gender and race, are normally incorporated as fixed effects. Therefore iddir and enumarator are random effects.

We use the regression line with standard algebraic notation to demonstrate a mixed-effects or multilevel model. Let Y represent the amount transferred and X the treatment group, $x_i = \{Trt1, Trt2, Trt3\}$. Assuming we have only one iddir, the regression line representing the relationship between a farmer's transfer and treatment group is:

$$y_i = \mu + \beta x_i + e_i, \tag{4.3}$$

where μ is the overall mean, β is the coefficient of the treatment and i is the index of farmers. It takes values 1 up to the number of farmers in that iddir ($i = 1, 2, \ldots, n_i$) and e_i is an error term or the residual, which represents the deviation of farmer $i's$ transfer from the mean level of transfer in i's treatment. With only one iddir, the variance of the e_i*'s* is the level 1 or farmer level variation. (4.3) can be rewritten as

$$\hat{y}_i = \mu + \beta x_i.$$

In a multilevel case, where we have more than one iddir and the iddirs are assumed to represent a random sample from a population of iddirs, the regression line is

$$\hat{y}_{ic} = \mu_c + \beta x_{ic}, \tag{4.4}$$

where c is an index of iddirs. It takes values 1 up to the number of iddirs in the sample, hence in our study, $c = 1, \ldots, 16$. The two subscripts indicate that the transfer y_{ic} varies from farmer to farmer within an iddir, similarly to x_{ic} the treatment group. Meanwhile, μ_c, with only one subscript is the same for all farmers in the same iddir but varies across iddirs. It is the average for each iddir, but can be broken down into a global average, denoted μ and the deviation of an iddir from the overall average, denoted u_c. Thus

$$\mu_c = \mu + u_c.$$

Therefore

$$\hat{y}_{ic} = \mu + \beta x_{ic} + u_c, \tag{4.5}$$

where μ is the overall average transfer and u_c is the deviation of the c^{th} iddir from the global mean. The variance of the u_c*'s* is the level 2 or iddir level variation.

In our multilevel case, the iddirs and enumerators are crossed factors at level two. Enumerators are assumed to be a random sample of enumerators, so the regression line becomes

$$y_{i(c,k)} = \mu + \beta x_{ic} + u_c + w_k + e_{i(c,k)}, \tag{4.6}$$

where k is an index of enumerators, $(k = 1, \ldots, 10)$ and u_c, w_k and $e_{i(c,k)}$ are assumed to independent, identical and normally distributed with mean zero and respective variances σ_u^2, σ_w^2 and σ_e^2, which are the *random* effects of our multilevel or mixed-effects model.

The three subscripts indicate that the transfers vary from farmer to farmer within an iddir-enumerator combination at level 2. It is assumed that training enumerators on the games and conducting a pilot study prior to conducting the experiments ensured minimal or inconsequential treatment variation due to enumerators. Therefore the treatment effect is assumed to vary from farmer to farmer within iddirs. Thus w_k is the deviation of the k^{th} enumerator from the global mean. The level 2 variation is a combination of the variances σ_u^2 and σ_w^2. Hence LMMs (4.5) and (4.6) are also known as a *variance components* models [?].

We adopted the practice of representing random parameters by Roman letters and constant parameters by Greek letters to distinguish Linear Models (LMs) from Linear Mixed-effects Models (LMMs) [?]. Hence, (4.4) is a LM, where the parameters μ and β are *constants* estimated from the data. Whereas (4.5) and (4.6) are LMMs which treat some of the parameters as *random*. We do not suppose that these "random variables" arise from a random process, rather, these can be understood as categorical covariates whose levels constitute a random sample from the set of all possible levels of that factor.

The normality assumption in both LMs and LMMs means that the expected value of the response variable is a linear combination of unknown parameters to be estimated from the data. However, as explained in Section (4.1.1) on GLMs, where the data cannot be assumed to be normally distributed, the mean cannot be assumed to be a linear combination of parameters, we therefore *generalize*, such that not the mean but some function of the mean can be assumed to be a linear combination of parameters estimated from the data. If the observations are count data, which follow a Poisson distribution with mean λ, a standard approach is to model $\log \lambda$ as a linear combination of the explanatory variables. If the parameters are ordinal data, which follow a multinomial distribution, the standard approach is to model the log odds as a linear combination of predictors. If the parameters are all constant then the result is a Generalized Linear Model (GLM), but if the parameters are mixed, some random and some constant, then the result is a multilevel model or Generalized Linear Mixed-effects Model (GLMM).

Notwithstanding, caution is necessary when level 1 units per level 2 units, that is the number of farmers within each iddir is small as the estimates will not be approximately normal even when a model is fitted correctly [**?**]. It is also important to note that if the variance components are estimated to be 0 it does not imply an absence of between-group variability but that the level of variability does not warrant incorporating random effects in the model. In which case, the mixed-effects model degenerates into a linear model. Therefore, a degenerate model is a mixed effects model in which the estimates of the variance components turn out to be zero, such that the mixed effects model degenerates to a linear model. If the mixed effects model is degenerate, Restricted Maximum Likelihood (REML) of variance components in a mixed effects model coincide with the variance estimates used in linear models. Due to this, the REML criterion is preferable to the ML estimates for mixed-effects models [**?**].

However, the advantage of using the Maximum Likelihood (ML) estimates is that the likelihood ratio test can be used to compare two models. The REML criteria cannot be used to compare nested models, that is models such that one is a reduced version of another, with different fixed-effects structures. This is because REML estimates the random effects by considering linear combinations of the data that remove the fixed effects, hence different fixed effects structures result in different likelihoods of the two models, making them incomparable. Moreover the likelihood test is only an approximation, and one of its most critical assumptions is that the parameters under the null are not on the boundary of the parameter space. Furthermore, when testing the fixed effects, the p-values generated by the LR test tend to be small, sometimes overstating the importance of some effects. Meanwhile when testing the random effects, the test tends to be conservative, that is the p-values tend to be larger than they should be and the true p-value will be lower. Hence, observing a significant effect gives confidence of a truly significant effect, whereas a small but insignificant outcome may require further approximations with more accurate methods, like Bootstrap techniques [**?**].

Mixed-effects residuals

Mixed-effects models, for example (4.5), do not produce individual iddir residuals. However, the residuals can be estimated from the observed data and estimated parameters.

Given that the observed value of transfers is y_{ic}, the predicted value is $\hat{y}_{ic}$, the deviation is $R_{ic} = y_{ic} - \hat{y}_{ic}$ and the true residual for each iddir is the mean of the R_{ic} for all i's in that iddir, which is denoted $R_{.c}$. But to obtain r_c, an estimate of iddir c's residual, the true residual is multiplied by a *shrinkage factor* due to correlation within the iddir, thus

$$\hat{r}_c = \left(\frac{\sigma_u^2}{\sigma_u^2 + \sigma_e^2/n}\right) * R_{.c}.$$

The shrinkage factor is always less than or equal to 1, thus the estimated residual is always less or equal to the true residual and as sample size increases, shrinkage increases too [**?**].

Diagnostics procedures for mixed-effects models

As is the case with OLS, detection of outliers and data points with large influence is an important diagnostic procedure. However, in multilevel models this is complicated by the fact that a response variable may be an outlier at any number of levels in the hierarchical structure and with respect to any number of explanatory variables in the model. In our case, an outlier may be at the farmer level or at the level of the iddir and enumerators. Moreover, analysis of residuals is ineffective for checking influence because by definition, influential points force the regression line close to it hence highly influential data points may have very small residuals [**?**].

4.2.2 The probability distribution of a Mixed-effects Model

A mixed model incorporates 2 random variables, $\mathcal{Y}$, an n-dimensional vector of response variables and $\mathcal{B}$ a q-dimensional vector of random effects. The response variable is observed but conditional on the unobserved but realized $\mathcal{B} = \mathbf{b}$ [**?**].

Distributions of the random variables

The probability model specifies the unconditional distribution of $\mathcal{B}$ and the conditional distribution $(\mathcal{Y}|\mathcal{B} = \mathbf{b})$. These two distributions are independent.

The random effects are unobserved, therefore in most cases are assumed to have the simple distribution

$$\mathcal{B} \sim N(\mathbf{0}, \mathbf{\Sigma}_{\boldsymbol{\theta}}),$$

where $\mathbf{\Sigma}_{\boldsymbol{\theta}}$ is a parameterized positive semi-definite symmetric matrix and $\boldsymbol{\theta}$ is a vector of parameters of the variance components, which usually has a small dimension of 1, 2 or 3.

The values $\mathbf{b}$ change only the conditional mean of the conditional distribution $(\mathcal{Y}|\mathcal{B} = \mathbf{b})$, through the *linear predictor*

$$\mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{b}.$$

Where $\boldsymbol{\beta}$ is a p-dimensional vector of fixed-effects and $\mathbf{X}_{n\times p}$ and $\mathbf{Z}_{n\times q}$ are *model matrices*, which are determined by the model formula and the corresponding covariate values. Although $\mathbf{Z}$ can be a large matrix it is usually sparse, that is most elements are zero.

In a linear mixed effects model (LMM), both the conditional distribution $(\mathcal{Y}|\mathcal{B} = \mathbf{b})$ and the unconditional distribution $\mathcal{B}$ are assumed to be normally distributed, and the conditional mean is the linear predictor

$$\boldsymbol{\mu}_{Y|B} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{b}.$$

Therefore

$$\begin{aligned} (\mathcal{Y}|\mathcal{B} = \mathbf{b}) &\sim N(\mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{b}, \boldsymbol{\sigma}^2\mathbf{I}) \\ \mathcal{B} &\sim N(\mathbf{0}, \boldsymbol{\Sigma}_{\boldsymbol{\theta}}). \end{aligned} \tag{4.7}$$

The variance components parameter $\boldsymbol{\theta}$ determines the *relative covariance factor* Λ_θ, a $q \times q$ matrix, which in turn generates a symmetric $q \times q$ variance covariance matrix $\boldsymbol{\Sigma}_\theta$, which defines the distribution of $\mathcal{B}$

$$\boldsymbol{\Sigma}_\theta = \boldsymbol{\sigma}^2 \Lambda_\theta \Lambda_\theta^T,$$

where $\boldsymbol{\sigma}$ is a common scalar parameter incorporated into the variance-covariance matrices of both $\mathcal{Y}$ and $\mathcal{B} = \mathbf{b}$.

If the normality assumption cannot be assumed then the conditional distribution can follow a non-normal distribution. This is defined as a generalized linear mixed effects model (GLMM). A Bernoulli distribution is commonly chosen for a binary response variable, a Poisson for count variable and a multinomial distribution for categorical response variable.

4.3 Generalized Linear Mixed-effects Model

This section brings together the discussions of the previous sections on GLMs and LMMs, which culminates in GLMMs.

A generalized linear mixed-effects model (GLMM) incorporates coefficients in a linear predictor but allows the response variable to follow a non-normal distribution. Additionally, the linear predictor incorporates random-effects and fixed effects parameters [????].

The response variables from our experiment, which are given in Table 4.1 are interval variables ranging from 0 to 40. As shown in plots in Section 3.6, most are skewed and do not have a bell-shaped distribution. In addition, we expect the transfers to be correlated within an iddir and within the same enumerator. Therefore, initially a GLMM seemed a reasonable modelling approach. We assumed a Poisson distribution because this is the standard approach in modelling count data [??]. We fitted a GLMM with Pair-Type and Treatment fixed effects and crossed random enumerator and iddir effects. However, this attempt at fitting a GLMM resulted in over-dispersion, which meant that there was more variability in the data than the model could account for. When there is over-dispersion in a Poisson GLMM, a common statistical procedure is to fit using a negative binomial link function [?]. Once again the approach was unsatisfactory based on diagnostic plots used to check for the validity of the assumptions of generalized linear mixed modeling.

Table 4.2 presents the dispersion parameters for the GLMMs fitted. If the deviance is large relative to the degrees of freedom of the residuals then there is over-dispersion, which is relatively large for the Poisson GLMM under all 3 possibilities of insurance availability, uptake and rejection for both subsamples B (with no aggregate shock and Y (with an aggregate shock). As expected, fitting the GLMM with a negative binomial link function corrects for over-dispersion. Hence, the dispersion parameters in row 5 are all less than 1. However, based on a Likelihood Ratio or Chi-squared test of goodness of fit, we reject

		Bavail	Yavail	Btake	Ytake	Breject	Yreject
	deviance	4029.2	883.0	4312.9	953.6	4507.5	1019.9
Poisson GLMM	df.resid	425	104	425	106	423	104
	dispersion	4.21	2.93	5.81	4.41	5.86	4.10
	Chi-sq p-value	0	0	0	<0	<0	<0
Neg Bin GLMM	dispersion	0.73	0.85	0.73	0.57	0.67	0.66
	Chi-sq p-value	0	0	0	<0	<0	<0

Table 4.2: GLMM diagnostics parameters

the null hypothesis that the model fit holds against the alternative that an empty model holds [**??**] for both the Negative Binomial and Poisson GLMM. The p-values of the test are given in rows 4 and 6 for the Poisson and Negative Binomial GLMM respectively. That the models poorly fit the data is further substantiated by the residuals' diagnostic plots, which we now turn to.

Plots of residuals against fitted values are used to check the assumption of independence of errors and responses. Figures 4.1, 4.2 and 4.3 are for models fitted to subsample B, with no aggregate shock, under insurance availability, take-up and rejection, respectively. Figures 4.4, 4.5 and 4.6 are for models fitted to subsample Y, with an aggregate shock.

QQ plots are used to check the assumption of normality of the residuals. Figure 4.7, of the Poisson GLMM residuals shows that the residuals violate this assumption. Figure 4.14 also suggests that the residuals of the negative binomial GLMM are non-normal and therefore the model is inadequate.

In conclusion, a Poisson GLMM proved unsatisfactory on account of overdispersion, low Chi-squared p-values and unsatisfactory residual diagnostic plots. Even after using a negative binomial to account for overdispersion, we found the models ill-fitting and unable to

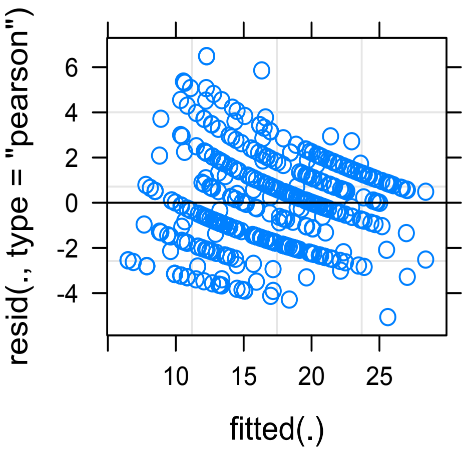

Figure 4.1: Poisson GLMM - subsample B under insurance availabilty. The residuals are heteroscedastic. As the fitted values increase, the residuals initially fan out then tapper off at high levels of fitted values, particularly in the positive range of residuals.

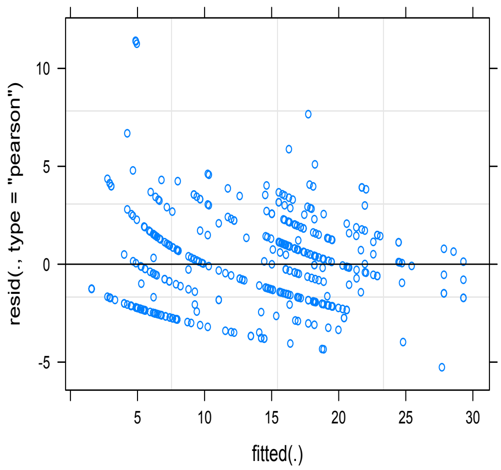

Figure 4.2: Poisson GLMM - subsample B under insurance takeup.

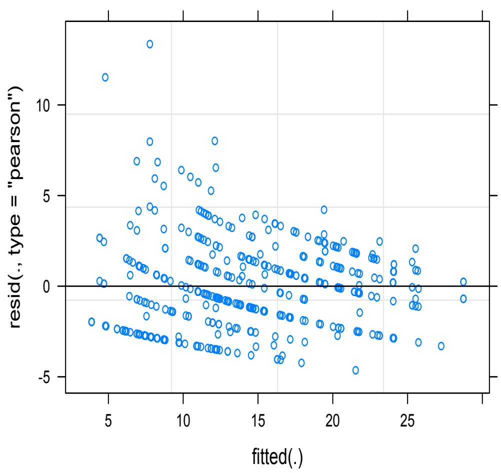

Figure 4.3: Poisson GLMM - subsample B under insurance rejection. The residuals are heteroscedastic due to tappering off at high levels of fitted values.

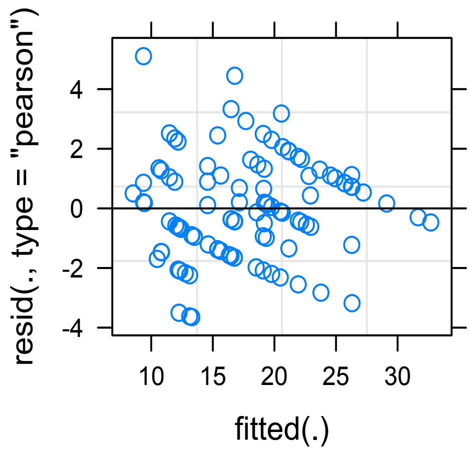

Figure 4.4: Poisson GLMM - subsample Y under insurance availabilty. The residuals are heteroscedastic. As the fitted values increase the residuals initially fan out then tapper off at high levels of fitted values, particularly in the positive range of residuals.

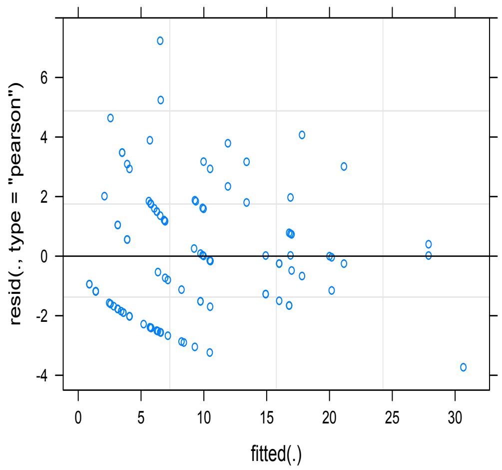

Figure 4.5: Poisson GLMM - subsample Y under insurance takeup. The residuals are heteroscedastic, they tapper off at high levels of fitted values.

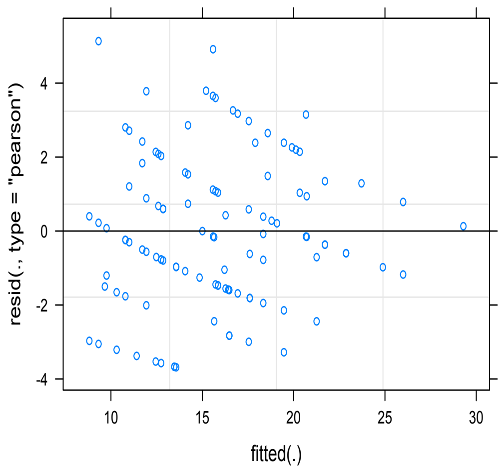

Figure 4.6: Poisson GLMM - subsample Y under insurance rejection. The residuals are heteroscedasticdue to the tappering off at high levels of fitted values.

capture the multimodality of our data. As a result of the empirical findings, we converted our response variables from interval to ordinal data by grouping the transfers into four groups, based on the inherent clustering of transfers at 0, 10, and 20 ETB. We elaborate on this approach in the next chapter.

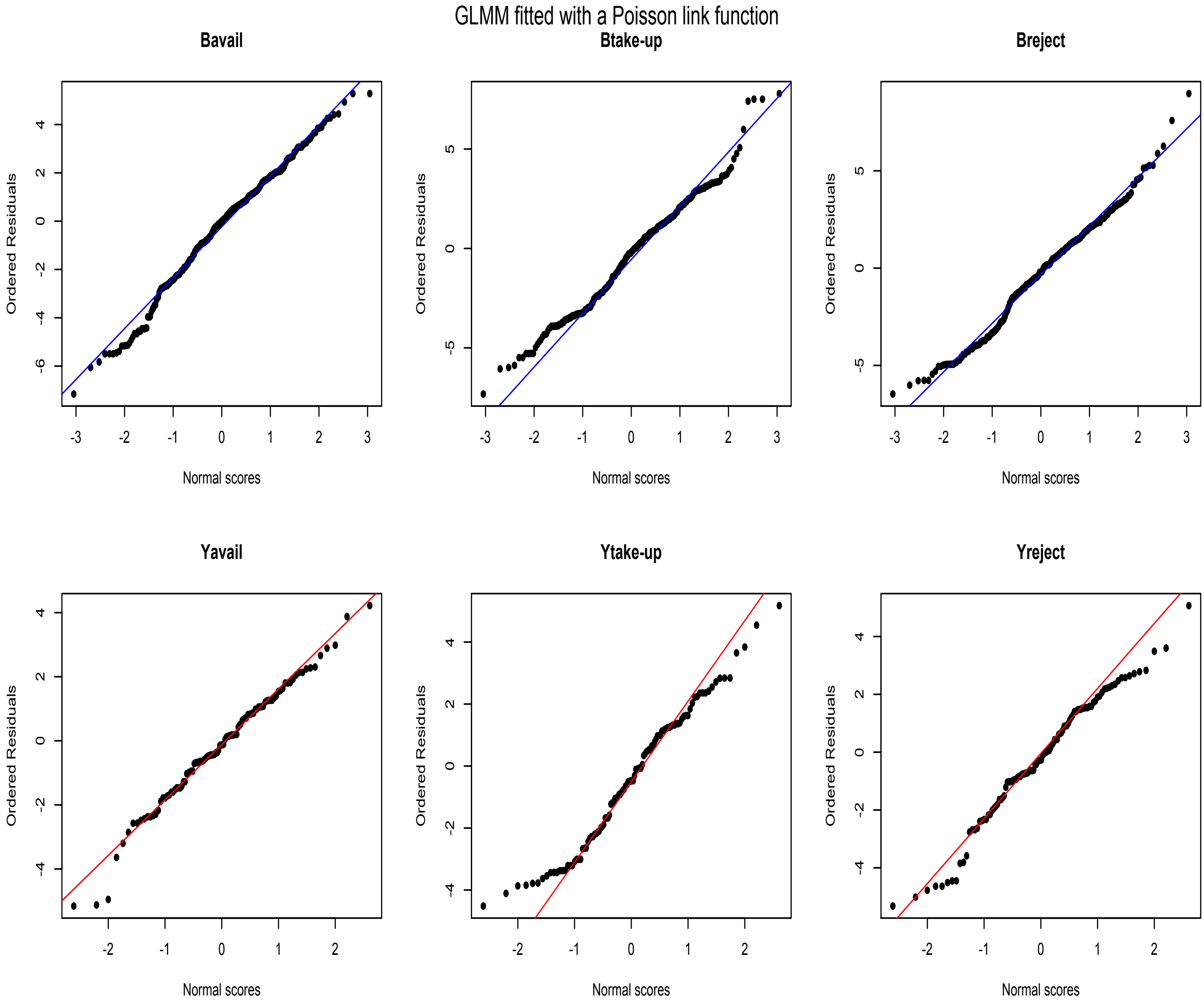

Figure 4.7: Poisson GLMM : Checking the assumption of normality of residuals.

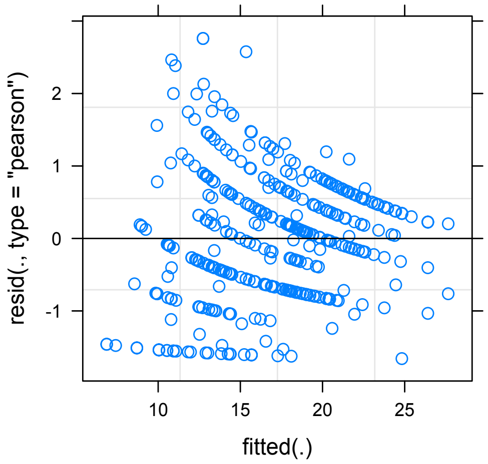

Figure 4.8: Negative Binomial GLMM - subsample B under insurance availability. The residuals are heteroscedastic. As the fitted values increase the residuals initially fan out then there residuals tapper off at high levels of fitted values, particularly in the positive range of residuals.

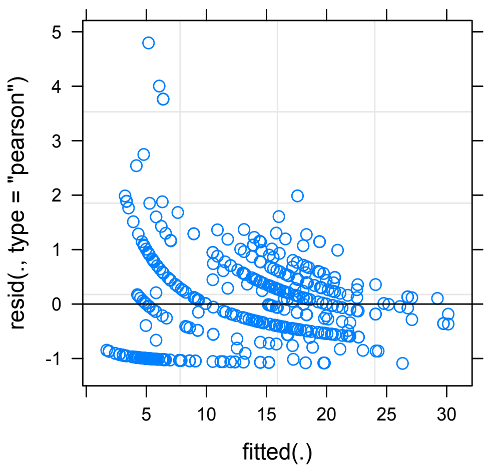

Figure 4.9: Negative Binomial GLMM - subsample B under insurance take-up

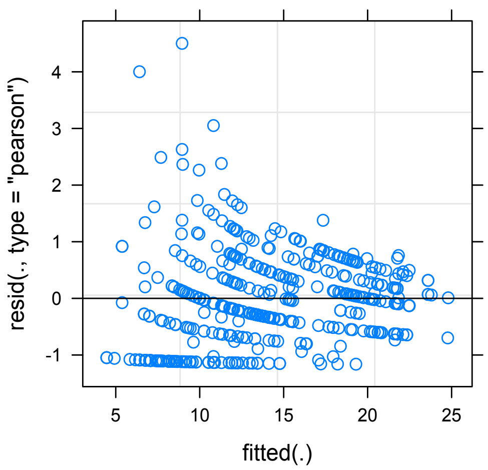

Figure 4.10: Negative Binomial GLMM - subsample B under insurance rejection

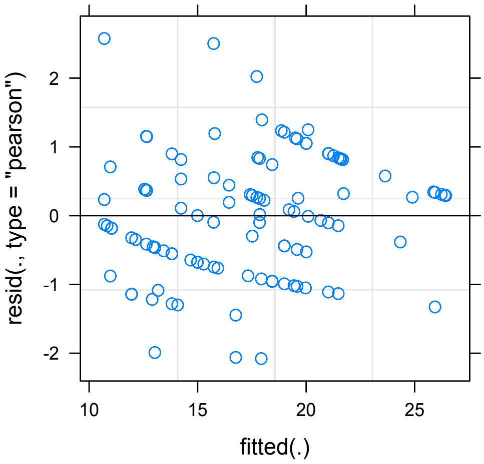

Figure 4.11: Negative Binomial GLMM - subsample Y under insurance availability.

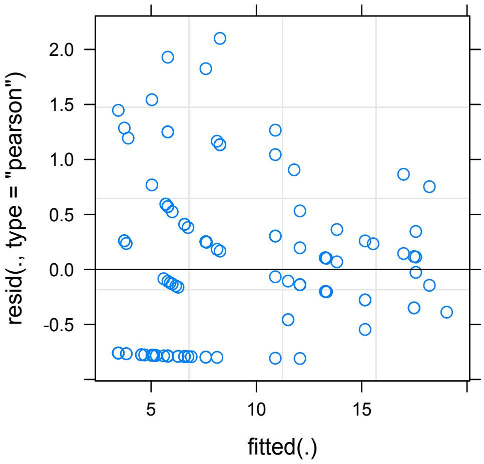

Figure 4.12: Negative Binomial GLMM - subsample Y under insurance take-up.

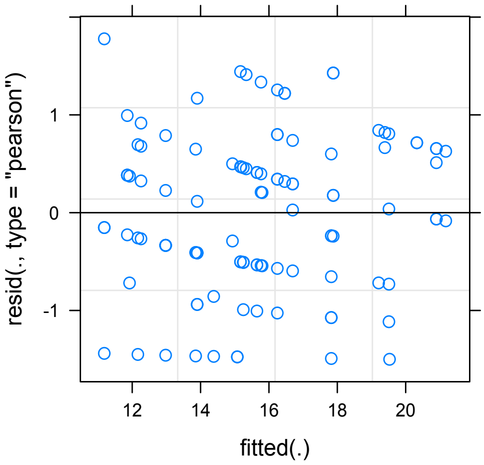

Figure 4.13: Negative Binomial GLMM - subsample Y under insurance rejection.

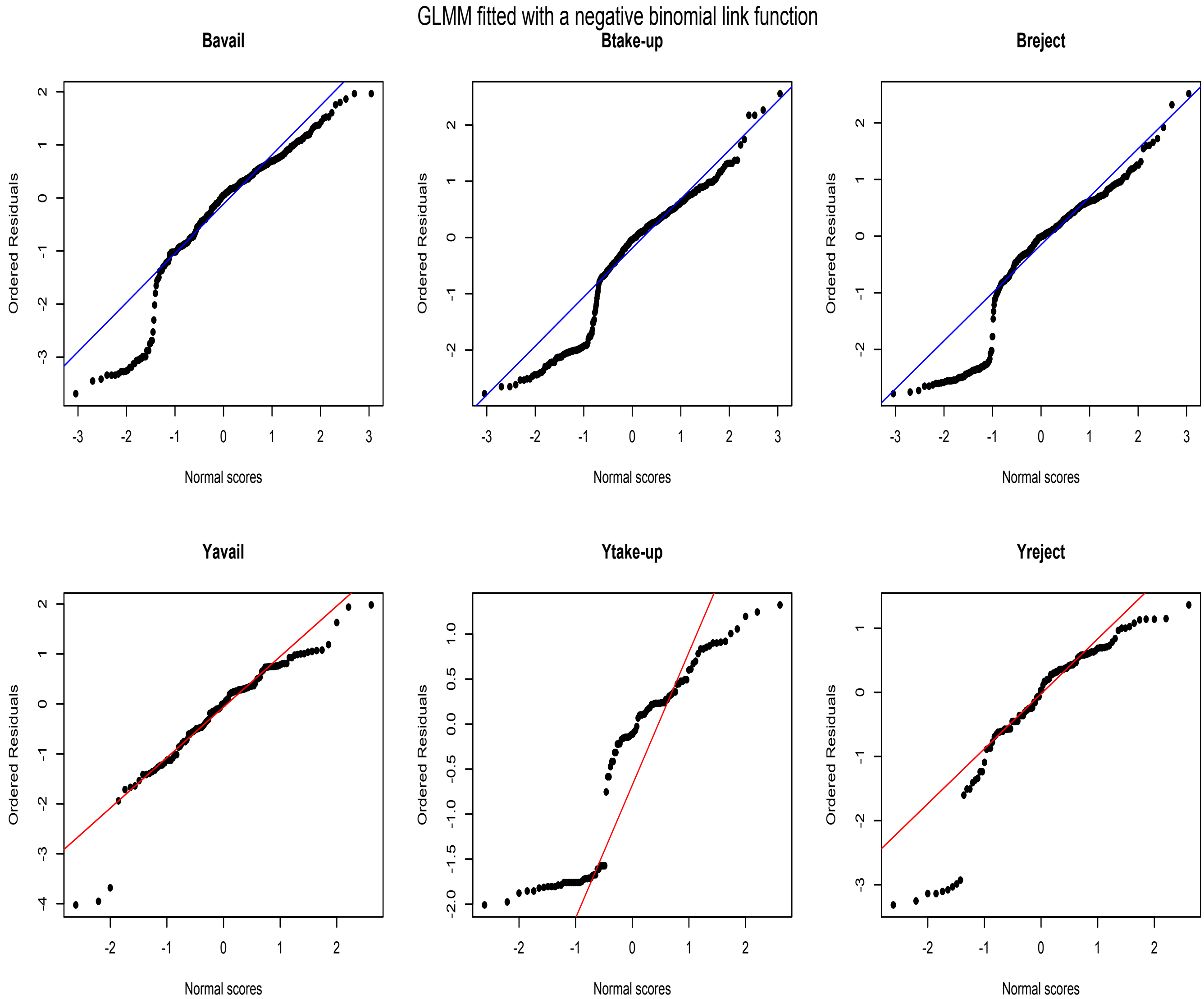

Figure 4.14: Negative Binomial GLMM : Checking the assumption of normality of residuals.

Chapter 5

Categorical response data

Our interval variables have a small number of distinct values, namely 0, 10, 20, and 30. Based on this as well as the multi-modality in the histogram of the transfers, we grouped the transfers into 4 categories (0, 1-10, 11-20 and 20+). This approach allows us to account for this underlying distribution of the transfers which a GLMM approach failed to do in the previous chapter. We begin by describing the multinomial distribution, thereafter, we discuss the ordinal logistic regression model. In Section 5.2 we specify the Cumulative Link Mixed-effects Model (CLMM) that we fitted to our data using the *"clmm"* function in the *"ordinal"* package in R [**??**]. We conclude by presenting the results of the CLMMs.

5.1 Multinomial models

GLMs with a multinomial response variable can be modelled in either of two ways. One entails expanding the logistic regression from binary response to nominal or ordinal responses for unordered and ordered responses respectively. An alternative procedure is log-linear modelling. In the former, one of the variables in the data is treated as a response variable, while the others are predictors, whereas in the latter, all variables are treated alike. Hence, log-linear modelling is ideal for cross-sectional studies with no clear response variable and complex multiway interactions [**?**].

The procedure of generalizing a binary logistic model to a multinomial logistic model is more appropriate for our experimental design and for the purposes of answering the research question, hence we limit the discussion to multinomial logistic regression.

5.1.1 Multinomial distribution

Suppose Y is a random variable with J categories such that $j = 1, 2, \ldots, J$ and X is a predictor variable. Let $\pi_j(x)$ denote the probability that Y lies in category j, so $P(Y = j \,|\, x) = \pi_j(x)$ and $\pi_1(x) + \pi_2(x) + \cdots + \pi_J(x) = 1$.

If there are n independent observations of Y then y_1 denotes the number of outcomes in the first category, y_2 denotes the number of outcomes in the second category and subsequently, y_j denotes the number of outcomes in the j^{th} category, such that $\sum\limits_{j=1}^{J} y_j = n$. The responses can be presented as a column vector

$$\mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_J \end{pmatrix}.$$

These counts at the J categories are multinomial variates, with multinomial distribution:

$$f(\boldsymbol{y}) = \frac{n!}{y_1!\, y_2! \ldots y_J!} \pi_1^{y_1} \cdot \pi_2^{y_2} \ldots \pi_J^{y_J}, \tag{5.1}$$

denoted $\mathcal{M}(n, \pi_1, \pi_2, \ldots, \pi_J)$.

Although the multinomial distributions do not form an exponential family, as they do not fit the general exponential form given in (4.1), they can be related to the Poisson, which then allows fitting a generalized linear model.

To see how a multinomial and Poisson can be related, suppose $Y_1, Y_2, \ldots, Y_J$ are independent random variables which follow a Poi(λ_j), their joint probability distribution is given by

$$f(\boldsymbol{y}) = \prod_{j=1}^{J} \frac{\lambda_j^{y_j} e^{-\lambda_j}}{y_j!}. \tag{5.2}$$

If $n = Y_1 + Y_2 + \cdots + Y_J$, then $n \sim Poi(\lambda_1 + \lambda_2 + \cdots + \lambda_J)$. Therefore given $N = n$, the distribution of $\boldsymbol{Y}$ is of the form

$$f(\boldsymbol{y}|n) = \left(\prod_{j=1}^{J} \frac{\lambda_j^{y_j} e^{-\lambda_j}}{y_j!} \right) \Bigg/ \frac{(\lambda_1 + \lambda_2 + \cdots + \lambda_J)^n e^{-(\lambda_1 + \lambda_2 + \cdots + \lambda_J)}}{n!}. \tag{5.3}$$

Now let $\pi_j = \frac{\lambda_j}{\sum_k \lambda_k}$ for $j = 1, 2, \ldots, J$ with $\sum_{j=1}^{J} \pi_j = 1$, then (5.3) simplifies to

$$f(\boldsymbol{y}|n) = \left(\frac{\lambda_1}{\sum \lambda_k} \right)^{y_1} \ldots \left(\frac{\lambda_J}{\sum \lambda_k} \right)^{y_J} \frac{n!}{y_1!\, y_2! \ldots y_J!}. \tag{5.4}$$

Therefore, a multinomial distribution can be regarded as a joint distribution of Poisson random variables conditional on their sum. This allows the use of GLMs for modelling multinomial responses. If there is no natural order of the response categories, the data is said to be nominal and the resulting model is a nominal logistic regression model. However, if the categories of a response variable are naturally ordered, then the data is ordinal. In this case, an ordinal logistic regression model is fitted, which we turn to next [?].

5.1.2 Ordinal logistic regression

The most common form of an ordinal logistic regression model is the cumulative logit model. We discuss this here as well as its specialized form, the proportional odds model.

Cumulative link model

Cumulative logits are log odds of cumulative probabilities. Since the responses are ordinal, a cumulative probability that the response is in less than a certain category has meaning. Assuming the Y is an ordinal response with J categories, the cumulative logit model uses the natural ordering to form cumulative probabilities,

$$P(Y \leq j|\mathbf{x}) = \pi_1(x) + \pi_2(x) + \cdots + \pi_j(x), \qquad j = 1, 2, \ldots, J.$$

From which cumulative logits can be defined as

$$\begin{aligned} \text{logit}[P(Y \leq j|\boldsymbol{x})] &= \log \frac{\pi_1(\boldsymbol{x}) + \cdots + \pi_j(\boldsymbol{x})}{\pi_{j+1}(\boldsymbol{x}) + \cdots + \pi_J(\boldsymbol{x})} \\ &= \log \frac{P(Y \leq j|\boldsymbol{x})}{1 - P(Y \leq j|\boldsymbol{x})} \\ &= \mathbf{x}_j^T \boldsymbol{\beta}_j \qquad\qquad j = 1, 2, \ldots, J. \end{aligned} \tag{5.5}$$

Hence the logit link function $G^{-1}(\mu) = \log[\mu/(1-\mu)]$ is the inverse of the standard logistic cdf.

Cumulative logit models simultaneously fit all (J–1) logit models for all the J response categories. Probabilities for the individual categories are found by subtracting adjacent cumulative probabilities.

$$P(Y = j|\boldsymbol{x}) = P(Y \leq j|\boldsymbol{x}) - P(Y \leq j-1|\boldsymbol{x})$$

Next, we discuss the most common case of the cumulative logit model, the proportional odds model, which assumes the same covariate effects for each logit, but different intercepts.

Proportional odds model

Suppose the linear predictor $\boldsymbol{x}_j^T\boldsymbol{\beta}_j$ has an intercept which depends on the category j, that is, β_{0j} but the other explanatory variables are independent of the response categories, then (5.5) takes the form

$$\begin{aligned}
\operatorname{logit}[P(Y \leq j|\boldsymbol{x})] &= \log \frac{\pi_1(\boldsymbol{x}) + \cdots + \pi_j(\boldsymbol{x})}{\pi_{j+1}(\boldsymbol{x}) + \cdots + \pi_J(\boldsymbol{x})} \\
&= \beta_{0j} + \beta_1 x_1 + \cdots + \beta_{p-1}x_{p-1} \\
&= \alpha_j + \beta_1 x_1 + \cdots + \beta_{p-1}x_{p-1},
\end{aligned} \tag{5.6}$$

where $\alpha_j = \beta_{0j}$.

Each cumulative logit has its own intercept α_j and because the cdf is an increasing function of j and the logit is an increasing function of the cdf the α_j are increasing in j.

The proportional odds model assumes that the effects of the covariates are the same for all categories. For any fixed j, for a single continuous predictor the response curve is a logistic regression or sigmoidal curve. The curves for all j categories of the same response have the same shape because they increase or decrease at the same rate. However, because they have different intercepts, they are horizontally displaced from one another [**?**].

According to the *cumulative odds ratio*

$$\begin{aligned}\text{logit}[P(Y \leq j|\boldsymbol{x}_1)] - \text{logit}[P(Y \leq j|\boldsymbol{x}_2)] &= \log \frac{P(Y \leq j|\boldsymbol{x}_1)/P(Y > j|\boldsymbol{x}_1)}{P(Y \leq j|\boldsymbol{x}_2)/P(Y > j|\boldsymbol{x}_2)} \\ &= \boldsymbol{\beta}^T(\boldsymbol{x}_1 - \boldsymbol{x}_2)\end{aligned}$$

the odds of making a response $\leq j$ at $\boldsymbol{x} = \boldsymbol{x}_1$ are $e^{\boldsymbol{\beta}^T(\boldsymbol{x}_1 - \boldsymbol{x}_2)}$ times the odds at $\boldsymbol{x} = \boldsymbol{x}_2$. This implies that the log cumulative odds ratio is proportional to the distance between $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$. In the proportional odds model, the proportionality is constant in each logit [**?**].

The proportional odds model is motivated by the assumption that there exists a continuous latent variable Y^* which drives the response process Y. The procedure of fitting an ordinal model entails identifying cut-points or thresholds $-\infty = \alpha_0 < \alpha_1 < \cdots < \alpha_J = \infty$ for the latent variable such that if the latent variable falls in the j^{th} interval, then the observed response y lies in category j:

$$\text{if } \alpha_{j-1} < y^* < \alpha_j, \qquad \text{then } y = j.$$

If the values of the latent variable vary around some level, denoted by some location parameter η that depends on the predictors $\boldsymbol{x}$ through $\eta(\boldsymbol{x}) = \boldsymbol{\beta}^T(\boldsymbol{x})$ and suppose Y^* has cdf $G(y^* - \eta)$, then

$$P(Y \leq j\,|\mathbf{x}) = P(Y^* \leq \alpha_j\,|\boldsymbol{x}) = G(\alpha_j - \boldsymbol{\beta}^T(\boldsymbol{x})).$$

This implies that the inverse cdf of the latent variable Y^* is the appropriate link function G^{-1} for the ordinal response Y. If G is the logistic function, $G(x) = e^x/(1 - e^x)$, then

$$P(Y \leq j\,|\boldsymbol{x}) = \frac{\exp(\alpha_j - \boldsymbol{\beta}^T(\boldsymbol{x}))}{1 + \exp(\alpha_j - \boldsymbol{\beta}^T(\boldsymbol{x}))},$$

and we have a logit model for the cumulative probabilities. Hence, the choice of a link variable depends on the distribution of the latent variable. If the latent variable is assumed to follow a normal distribution, we have a probit model for the cumulative probabilities. If the latent variable is assumed to follow a Gumbel distribution, the Type I extreme-value distribution, then we have a complementary log-log model for the cumulative probabilities [**?**].

In a proportional odds model, β does not depend on j, which means the predictors have a uniform effect on all the response categories. The same parameters $\boldsymbol{\beta}$ for the effects occur regardless of the threshold points. If the categories are collapsed, the parameter estimates do not change. The cut points and the explanatory variables are independent [**??**].

Goodness of fit measures

Various measures can be used to estimate goodness of fit. The **Pearson chi-squared residuals**

$$r_i = \frac{o_i - \hat{e}_i}{\sqrt{\hat{e}_i}}$$

where $i = 1, 2, \ldots, N$, with $N = J\times$ the number of distinct covariate patterns and o_i is the observed frequency and $\hat{e}_i$ the expected frequency produce a **Chi-squared statistic**

$$\mathcal{X}^2 = \sum_{i=1}^{N} r_i^2$$

In addition, we can use the **Deviance**, which is twice the difference between the maximum value of the log likelihood function for the fitted model ($\mathcal{L}_m$) and the maximum value of the log likelihood function for the saturated ($\mathcal{L}_s$) model,

$$\mathcal{D} = 2(\mathcal{L}_s - \mathcal{L}_m) \sim \mathcal{X}^2(N - p)$$

Or the **Likelihood ratio chi-squared statistic**, $\mathcal{LR}$ which is twice the difference between the maximum value of the log likelihood function of the fitted model ($\mathcal{L}_m$) and the maximum value of the log likelihood function of the null ($\mathcal{L}_0$) model,

$$\mathcal{LR} = 2(\mathcal{L}_m - \mathcal{L}_0) \sim \mathcal{X}^2(p - (J - 1))$$

because the minimal model will have one parameter for each logit in the model.

We can also use the **Pseudo coefficient of determination** $\mathcal{R}^2$

$$\mathcal{R}^2 = \frac{\mathcal{L}_0 - \mathcal{L}_m}{\mathcal{L}_0},$$

as well as the Akaike Information Criterion (AIC) to compare non-nested models.

5.1.3 Fitting a CLMM - parameter estimation

Similarly, to a GLMM, the likelihood function of a CLMM does not have a closed form, hence model fitting is complex. Although computationally intensive, it can be approximated by numerical methods. Once approximated, the likelihood can then be maximized by standard algorithms like the Newton-Raphson to produce ML estimates $\hat{\boldsymbol{\beta}}$ and $\hat{\boldsymbol{\Sigma}}$. Standard errors are then obtained by inverting the observed information matrix [**?**].

As in a GLMM, a CLMM is a two-stage model. In the first stage, the observations are assumed to follow a GLM, conditional on the random effects $\{\boldsymbol{u}_i\}$. Hence, conditional on the random effects, the observations y_{ic} are assumed to be independent and from an exponential family, with expected value μ_{ic} linked to a linear combination of the predictors, that is, the linear predictor $\boldsymbol{\eta}$, through a link function $g()$, which in a CLMM is a cumulative log odds function:

$$g(\mu_{ic}) = \boldsymbol{x}_{ic}^T\boldsymbol{\beta} + \boldsymbol{z}_{ic}^T\boldsymbol{u}_i \tag{5.7}$$

Where $\boldsymbol{u}_i$ denotes the vector of random effects.

In the second stage, the random effects $\boldsymbol{u}_i$ are assumed to be independent $N(\mathbf{0}, \boldsymbol{\Sigma})$. Letting $\boldsymbol{y}$ denote the vector of observations, the conditional mass function of $\boldsymbol{y}$ given $\boldsymbol{u}$ is $f(\boldsymbol{y}\,|\boldsymbol{u};\boldsymbol{\beta})$ and let $f(\boldsymbol{u};\boldsymbol{\beta})$ denote the normal probability density for $\boldsymbol{u}$. Hence the marginal likelihood function $\ell(\boldsymbol{\beta},\boldsymbol{\Sigma};\boldsymbol{y})$ for a GLMM is the probability mass function $f(\boldsymbol{y};\boldsymbol{\beta},\boldsymbol{\Sigma})$,

$$\ell(\boldsymbol{\beta},\boldsymbol{\Sigma};\boldsymbol{y}) = f(\boldsymbol{y};\boldsymbol{\beta},\boldsymbol{\Sigma}) = \int f(\boldsymbol{y}\,|\boldsymbol{u};\boldsymbol{\beta})f(\boldsymbol{u};\boldsymbol{\beta})\,d\boldsymbol{u} \tag{5.8}$$

Hence, the Likelihood function results from integrating out the random effects from the joint distribution of $\boldsymbol{y}$ and $\boldsymbol{u}$.

The marginal likelihood for the logistic normal random effects model is

$$\ell(\boldsymbol{\beta},\sigma^2;\boldsymbol{y}) = \prod_i \left(\int_{-\infty}^{\infty} \prod_c \left[\frac{\exp\left(\boldsymbol{x}_{ic}^T\boldsymbol{\beta}+\boldsymbol{u_i}\right)}{1+\exp\left(\boldsymbol{x}_{ic}^T\boldsymbol{\beta}+\boldsymbol{u_i}\right)}\right]^{y_{ic}} \left[\frac{1}{1+\exp\left(\boldsymbol{x}_{ic}^T\boldsymbol{\beta}+\boldsymbol{u_C}\right)}\right]^{1-y_{ic}} f(u_i;\sigma^2)du_i \right) \tag{5.9}$$

There are standard numerical integration methods for approximating the likelihood function when the dimension of the random effects structure is small. These include the Gauss-Hermite quadrature method, which approximates the integral of a function by multiplying it by a normal density function and producing a finite weighted sum evaluated at certain specified points known as quadrature points. The approximation improves with the number of quadrature points. For higher dimensions of the random effects Monte Carlo methods are more computationally feasible than numerical integration methods. We do not detail these more complex methods as the low dimensionality of our random effects does not call for this approach.

Alternatively, the Laplace and Penalized Quasi-likelihood methods can be used. The approximation maximises an analytical approximation of the likelihood function. From (5.8), the likelihood function is obtained by integrating out the the random effects from the joint distribution of $\boldsymbol{y}$ and $\boldsymbol{u}$. Now, the Laplace and penalized quasi-likelihood method take the exponential family representation of each component of the joint distribution and the integrand in (5.8) becomes an exponential function of $\boldsymbol{u}$.

The Laplace approximation then uses a second order Taylor series expansion of the exponent around a point $\tilde{u}$, at which the first order approximation equals zero, so $\tilde{\boldsymbol{u}} = E(\boldsymbol{u}|\boldsymbol{y})$. The exponential function approximating the integrand has a quadratic exponent in $(\boldsymbol{u} - \tilde{\boldsymbol{u}})$ which takes the form of a constant multiple of a multivariate normal density. Hence the approximate integral has a closed form and can be maximized with respect to $\boldsymbol{\beta}$ and $\boldsymbol{\Sigma}$.

A method of approximating the integral by Breslow and Clayton (1993) in [?] yields a function approximating the log likelihood, which is of the form

$$q(\boldsymbol{\beta}, \boldsymbol{y}) - (1/2)\tilde{\boldsymbol{u}}^T \boldsymbol{\Sigma}^{-1} \tilde{\boldsymbol{u}}$$

where $q(\boldsymbol{\beta}, \boldsymbol{y})$ resembles a GLM quasi-likelihood function conditional on $(\boldsymbol{u} = \tilde{\boldsymbol{u}})$. Thus in the above form the second term is a penalty for the quasi-likelihood, which increases as elements of $\tilde{\boldsymbol{u}}$ increase in absolute value. Hence the approach is called the penalized quasi-likelihood (PQL).

Although the Laplace and PQL methods are computationally simpler to use, unlike the more complex methods that allow parameter estimates to converge to the ML estimates, they do not yield ML estimates. PQL estimates are particularly poor when the true variance components are large and when the response variable is very far from normal [**?**].

5.2 Specification of the CLMM

A cumulative logit link mixed model of the probability that the i^{th} response in cluster c falls below the j^{th} category is of the general form

$$\text{logit}\left[P(Y_{ic} \leq j|\boldsymbol{u}_c)\right] = \boldsymbol{\alpha}_j + \boldsymbol{x}_{ic}^T\boldsymbol{\beta} + \boldsymbol{z}_{ic}^T\boldsymbol{u}_c \qquad j = 1, \ldots, J-1. \tag{5.10}$$

Therefore the hierarchical or clustered nature of the data is captured.

With categorical explanatory variables, there is a parameter for each category of each explanatory variable. However, one parameter of each explanatory variable will be redundant. Hence, to prevent this we either constrain the sum of the parameters of each variable to zero. Alternatively we can set the parameter of one of the categories or levels of an explanatory variable to be equal to zero and define the parameters of the rest of the categories as deviations from the category set to zero. We prefer the latter, as it allows for more intuitive comparative interpretation.

Now, to represent the relationship between a donor's transfer and the donor's treatment group and pair-type:

> Let *Trt* denote treatment, which is nominal with 3 categories, Trt1, Trt2 and Trt3. β_0, the parameter for Trt1, is set to zero. Thus the parameters β_1 and

β_2 are respective deviations of Trt2 and Trt3 from Trt1.

Let S denote Pair-Type, which is nominal variable with 2 categories. Let $s = 1$ denote same iddir and $s = 0$ different iddirs. We estimate one parameter for the pair-type effect, thus β^S is the coefficient measuring the deviation of same-iddir transfers from different-iddir transfers.

Let i be an index of farmers in each iddir, whereby $(i = 1, \ldots, n_i)$, with c an iddir index whereby $(c = 1, \ldots, 16)$ and k an enumerator index, whereby $(k = 1, \ldots, 10)$. Recall, that we selected 48 donor farmers from each of 16 iddirs, and randomly assigned each farmer to either of 10 enumerators during the experiment. Thus farmers are nested or clustered within iddirs which are crossed with enumerators.

Therefore, we fit a model for the cumulative probability that $\mathcal{T}_{i(c,k)}$, a transfer made by farmer i in iddir c enumerated by k falls in level j of the transfers, where

$$j = \left\{ \begin{array}{lc} 1 & \text{if transfer} = 0 \\ 2 & \text{if } \ 0 < \text{transfer} <= 10 \\ 3 & \text{if } \ 10 < \text{transfer} <= 20 \\ 4 & \text{if transfer} > 20 \end{array} \right\}.$$

Hence we fit a cumulative logit link mixed-effect model that the i^{th} farmer's transfer falls in the j^{th} category or below, with $j = 1, \ldots, 4$ an index of the response categories. This is specified as:

$$\text{logit}[P(\mathcal{T}_{i(c,k)} \leq j|u_c, w_k, e_{i(c,k)})\,] = \alpha_j + \beta_1\, I(Trt = Trt2) + \beta_2\, I(Trt = Trt3) + \beta^S + u_c + w_k + e_{i(c,k)} \tag{5.11}$$

with $u_c \sim N(0, \sigma_u^2)$, and $w_k \sim N(0, \sigma_w^2)$ and $e_{i(c,k)} \sim N(0, \sigma_e^2)$.

Thus α_j are the threshold parameters or cut-points. We assume that the predictors are independent of the response categories. We assume that the random iddir and enumerator effects are iid and independent of the farmer level residuals.

From (5.11), the log odds ratios are

$$\text{logit}\ [P(\mathcal{T}_{i(c,k)} \leq j\ |u_c, w_k, e_{i(c,k)})\,] = log \frac{P(\mathcal{T}_{i(c,k)} \leq j)}{P(\mathcal{T}_{i(j)c,k} > j)}, \tag{5.12}$$

letting $P(\mathcal{T}_{i(c,k)} = j) = \pi_j$

and the log odds of the respective categories are:

$$\log\left(\frac{\pi_1}{\pi_2 + \pi_3 + \pi_4}\right) = \alpha_1 + \beta_1\, I(Trt = Trt2) + \beta_2\, I(Trt = Trt3) + \beta^S + u_c + w_k + e_{i(c,k)},$$

$$\log\left(\frac{\pi_1 + \pi_2}{\pi_3 + \pi_4}\right) = \alpha_2 + \beta_1\, I(Trt = Trt2) + \beta_2\, I(Trt = Trt3) + \beta^S + u_c + w_k + e_{i(c,k)},$$

$$\log\left(\frac{\pi_1 + \pi_2 + \pi_3}{\pi_4}\right) = \alpha_3 + \beta_1\, I(Trt = Trt2) + \beta_2\, I(Trt = Trt3) + \beta^S + u_c + w_k + e_{i(c,k)},$$

and

$$\pi_4 = 1 - (\pi_1 + \pi_2 + \pi_3).$$

To illustrate, we calculate the estimated probability that a donor transfers more than 20 ETB. The formula takes the form

$$\text{logit}\ [P(\mathcal{T}_{i(c,k)} > 3] = 1 - \left(\frac{e^{\alpha_3 + \beta_1 I(Trt=Trt2) + \beta_2 I(Trt=Trt3) + \beta^S}}{1 + e^{\alpha_3 + \beta_1 I(X=1) + \beta_2 I(X=2) + \beta^S}} \right).$$

		T_{ij}^{avail}		$T_{ij}^{take-up}$		T_{ij}^{reject}	
		Trt2	Trt3	Trt2	Trt3	Trt2	Trt3
B	PTd	0.33	0.16	0.02	0.15	0.09	0.21
	PTs	0.38	0.19	0.02	0.18	0.22	0.10
Y	PTd	0.35	0.11	0.01	0.01	0.23	0.07
	PTs	0.47	0.17	0.02	0.02	0.39	0.15

Table 5.1: The estimated probability that a donor transfers more than 20 ETB under insurance availability, take-up and rejection (Trt2 is indemnity insurance, Trt3 is index insurance). Sample B does not experience an aggregate shock, sample Y has an aggregate shock.

5.3 Results and discussion

In the previous section we specified a proportional odds model or cumulative logit link mixed-effects model that we fit to the data, see (5.11). The results of using the control group with no insurance (Trt1) as the baseline category are displayed in Table 5.2, while Table 5.3 presents the same model with the indemnity treatment group (Trt2) as the baseline for a direct comparison of Trt2 and Trt3.

During the model building process we evaluated different models using the AIC criterion. We explored models with Nr.Adults (number of adults in the donor's household), Edu (donor's level of education) and Occ (donor's occupation) as these variables proved unbalanced across the treatment groups in Section 3.1. Notwithstanding the AIC criterion, at a p-value of 5% these predictors proved statistically insignificant under most of the treatment conditions. Moreover, our aim is to investigate Treatment and Pair-Type effects, therefore our discussion will dwell on models with only these two predictors. These models are reasonably good fits, based on the Likelihood ratio test. Furthermore, randomization justifies our decision to discard the more complex models and to focus our analysis on the simpler model with only Treatment and Pair-Type (PT) effects. Having settled on a model of choice, we thus used the more complex models with Nr.Adults and Education or Occupation for robustness checks.

In both Table 5.2, and Table 5.3, the coefficient β_1 measures the deviation of Trt2 from Trt1, while β_2 measures the deviation of Trt3 from Trt1 and β^S measures the deviation of transfers in same iddir pair-types from different iddir pair-types. The table also contains the estimated threshold parameters α_1, α_2 and α_3 for the first (0 ETB), second ($1-10$ ETB) and third ($11-20$ ETB) response categories respectively. Finally, the tables show σ_u^2 and σ_w^2, the estimates of the iddir and enumerator random effects. The tables present the results of fitting the same CLMM, which is specified in (5.11), to sample B and sample Y. Recall, B is the sample of farmers who did not experience an aggregate shock and sample Y consists of farmers who experienced an aggregate shock. We use three different response variables for the different possibilities of insurance availability, take-up and rejection. The headings of the columns indicate the respective sample and response variable used in the model summarised by that column. The first three columns are fit to sample B, while the latter three are fit to sample Y. Columns 1 and 4 use the response variable T_{ij}^{avail}, which was elicited by asking a donor how much he or she would transfer if his or her partner suffered a loss. Columns 2 and 5 use the response variable $T_{ij}^{take-up}$, which was elicited by asking a donor how much he or she would transfer if his or her partner took up the insurance and suffered a loss. Similarly, Columns 3 and 6 use the response variable T_{ij}^{reject}, which was elicited by asking a donor how much he or she would transfer if his or her partner did not take up the insurance and suffered a loss.

The tables also provide goodness of fit measures, namely, the $\mathcal{LR}$-statistic, the pseudo $\mathcal{R}^2$ and the AIC. The $\mathcal{LR}$-statistic for each model, which is shown in the 9th row compares the choice model to a null model and is asymptotically $\mathcal{X}^2(p-(J-1))$ under the null hypothesis that the model fit holds against the alternative that an empty model holds [**??**]. Since our models estimate 8 parameters with $J = 4$, the $\mathcal{LR} \sim \mathcal{X}^2(5)$. Comparing the $\mathcal{LR}$-statistics to 11.07, the chi-squared statistic with 5 degrees of freedom at 5% level of significance shows that Treatment and Pair-Type have predictive power and the models are reasonably good fits. However, the pseudo $\mathcal{R}^2$ suggests that at most, only 11.2% of the variation is explained by these predictors. This is not alarming in a social science experimental study which aims to isolate and explain one mechanism or source of variation rather than to explain most of the variation.

Finally, to get a more nuanced understanding of the effect of Pair-Type, we stratified the two samples, Y and B, according to pair-type. We then fit a CLMM similar to the CLMM specified in (5.11) but without a Pair-Type fixed effect. The results are displayed in Table 5.4. The first 3 columns are models fit to a subsample of farmers from the same iddir who did not experience an aggregate shock (BPTs). The second 3 are fit to a subsample of farmers from different iddirs who did not experience an aggregate shock

Effect	$B\,T_{ij}^{avail}$	$B\,T_{ij}^{take-up}$	$B\,T_{ij}^{reject}$	$Y\,T_{ij}^{avail}$	$Y\,T_{ij}^{take-up}$	$Y\,T_{ij}^{reject}$
β_1(Trt2)	0.324	-2.928^{***}	-0.484^{*}	1.186^{*}	-2.259^{***}	0.471
	(0.241)	(0.284)	(0.241)	(0.467)	(0.505)	(0.439)
β_2(Trt3)	-0.596^{**}	-0.661^{**}	-1.475^{***}	-0.279	-2.374^{***}	-0.845
	(0.224)	(0.229)	(0.234)	(0.611)	(0.370)	(0.394)
β^S(s)	0.207	0.217	0.093	0.498	0.472	0.791^{*}
	(0.186)	(0.188)	(0.186)	(0.384)	(0.391)	(0.367)
α_1	-2.947	-2.786	-2.649	-3.200	-2.324	-2.352
α_2	-0.458	-0.514	-0.462	0.140	0.035	-0.079
α_3	1.036	1.043	0.859	1.806	1.998	1.696
σ_u^2	0.370	0.409	0.389	0.322	0.767	0.00
σ_w^2	0.448	0.516	0.509	0.041	0.000	0.392
$\mathcal{LR}$	28.316	145.165	53.646	13.606	30.477	11.802
pseudo $\mathcal{R}^2$	0.027	0.126	0.048	0.042	0.112	0.053
AIC	1051.24	1020.35	1079.72	260.02	258.08	285.66
Log Likelihood	-517.62	-502.18	-531.86	-122.01	-121.04	-134.83
Num. Iddir	16	16	16	10	10	10
Num. Enum	10	10	10	10	10	10
Num. obs	431	431	429	110	110	110

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$

Table 5.2: CLMMs models: The effect of insurance (both indemnity (Trt2) and index (Trt3)) availability, take-up and rejection on transfers without aggregate shock (B) and with an aggregate shock (Y), with Trt1 as the baseline category

(BPTd). The subsequent 3 models are fit the subsample of farmers from the same iddir who had an aggregate shock (YPTs) and the last three are fit to farmers from different iddirs (YPTd). Again the headings indicate the response variable used in that model.

Effect	$B\,T_{ij}^{avail}$	$B\,T_{ij}^{take-up}$	$B\,T_{ij}^{reject}$	$Y\,T_{ij}^{avail}$	$Y\,T_{ij}^{take-up}$	$Y\,T_{ij}^{reject}$
β_1(Trt1)	-0.324	2.928^{***}	0.484^{*}	-1.186^{*}	2.259^{***}	-0.471
	(0.241)	(0.284)	(0.241)	(0.467)	(0.505)	(0.439)
β_2(Trt3)	-0.920^{***}	2.267^{***}	-0.991^{***}	-1.465^{**}	-0.115	-1.316^{*}
	(0.246)	(0.276)	(0.249)	(0.562)	(0.531)	(0.623)
β^S(s)	0.207	0.217	0.093	0.498	0.4724	0.791^{*}
	(0.186)	(0.188)	(0.186)	(0.384)	(0.391)	(0.367)
α_1	-3.271	0.142	-2.166	-4.386	-0.064	-2.823
α_2	-0.782	2.415	0.022	-1.046	2.295	-0.551
α_3	0.712	3.971	1.343	0.621	4.257	1.225
σ_u^2	0.370	0.409	0.389	0.322	0.767	0.000
σ_w^2	0.448	0.516	0.509	0.041	0.000	0.392
$\mathcal{LR}$	28.316	145.1645	53.646	13.606	30.477	11.802
pseudo $\mathcal{R}^2$	0.027	0.126	0.048	0.042	0.112	0.053
AIC	1051.24	1020.35	1079.72	260.02	258.08	285.66
Log Likelihood	-517.62	-502.18	-531.86	-122.01	-121.04	-134.83
Num. Iddir	16	16	16	10	10	10
Num. Enum	10	10	10	10	10	10
Num. obs	431	431	429	110	110	110

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$

Table 5.3: CLMMs models: The effect of insurance availability, take-up and rejection on transfers without aggregate shock (B) and with an aggregate shock (Y), with Trt2 as baseline category

	BPTs			BPTd			YPTs			YPTd		
Effect	T_{ij}^{avail}	$T_{ij}^{take-up}$	T_{ij}^{reject}	T_{ij}^{avail}	$T_{ij}^{take-up}$	T_{ij}^{reject}	T_{ij}^{avail}	$T_{ij}^{take-up}$	T_{ij}^{reject}	T_{ij}^{avail}	$T_{ij}^{take-up}$	T_{ij}^{reject}
β_1(Trt2)	0.324	-2.879***	-0.444	0.440	-3.112***	-0.481	0.935	-2.284**	0.323	1.881**	-1.621*	0.685
	(0.334)	(0.410)	(0.330)	(0.365)	(0.424)	(0.372)	(0.800)	(0.745)	(0.665)	(0.620)	(0.650)	(0.564)
β_2(Trt3)	-0.294	-0.516	-1.219***	-0.868**	-0.834*	-1.775***	1.744.	-3.118**	-2.115*	0.250	-2.026**	-0.049
	(0.32)	(0.343)	(0.331)	(0.330)	(0.335)	(0.346)	(1.04)	(0.957)	(1.058)	(0.66)	(0.739)	(0.651)
α_1	-2.98	-2.9833	-2.5804	-3.1460	-2.9869	-2.8853	-6.20	3.0744	-4.418	-2.7660	-2.0369	-1.9130
α_2	-0.49	-0.548	-0.455	-0.646	-0.766	-0.621	-1.93	0.660	-1.593	0.404	0.179	0.221
α_3	0.87	0.845	0.803	1.080	1.161	0.809	0.22	1.24	0.412	2.1370	2.531	2.127
σ_u^2	0.21	0.430	0.259	0.6233	0.792	0.000	0.66	0.853	0.000	0.000	0.000	0.000
σ_w^2	0.29	0.437	0.28	0.726	0.689	0.392	3.22	0.301	3.68	0.000	0.159	0.010
$\mathcal{LR}$	3.607	67.172	14.450	13.845	66.701	28.962	7.915	13.722	7.034	2.035	9.502	10.300
pseudo$\mathcal{R}^2$	0.007	0.114	0.025	0.027	0.123	0.055	0.068	0.103	0.054	0.076	0.072	0.014
AIC	557.06	534.31	582.70	508.50	490.46	285.66	122.83	133.80	137.44	138.84	134.82	154.21
Log Likelihood	-271.53	-260.16	-284.35	-247.25	-238.23	-134.83	-54.42	-59.90	-61.72	-62.42	-61.41	-70.11
Num. Iddir	16	16	16	16	10	10	10	10	10	10	10	10
Num. Enum	10	10	10	10	10	10	10	10	10	9	9	9
Num. obs	222	222	222	207	207	205	53	53	53	57	57	57

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$

Table 5.4: CLMMs: The effect of insurance availability, take-up and rejection on transfers without aggregate shock for same iddir pairs (BPTs) and different iddir pairs (BPTd) and with aggregate shock for same iddir pairs (YPTs) and different iddir pairs (YPTd)

5.3.1 The effect of Index Insurance

The coefficients β_2, in the second row of Table 5.2 measure the effect of index insurance. This coefficient falls away if a donor did not receive the index treatment (Trt3) as shown in (5.11). The estimated effects suggest that the cumulative probability, starting at the lower end of the transfers scale, decreases when index insurance is available, taken up or rejected, with or without an aggregate shock.

When there is no aggregate shock

When index insurance is available, transfers are significantly lower than transfers in the control. The estimated odds of a transfer are $e^{-0.5963} = 0.55$ times the odds of a transfer in the control. When index insurance is available, a recipient's net payoff will either be 10 ETB or 28 ETB. A recipient's need for a transfer will either be heightened or equal to the need in the control group, yet we find a significant reduction in transfers. The fact that donors do not offer compensation for basis risk suggests that transfers are motivated by more than just the needs of the recipient. Hence, we do not observe a warranted increase in transfers. Recall that we defined such a mutually beneficial relationship as supplementarity. Therefore, we find no evidence of supplementarity between index insurance and transfers in our non-reciprocal setting.

When a recipient takes up index insurance and there is no aggregate shock, he or she faces basis risk since a payout is not triggered. The recipient's net payoff is 10 ETB compared to the the net payoff of 28 ETB in the control. Thus basis risk exacerbates the need for a transfer, yet transfers decrease compared to the control group. The estimated odds of a transfer are $e^{-0.661} = 0.52$ times the odds in the control. Donors are 48%

less likely to make a transfer in each response category. This finding is a deviation our predictions in Section (1.3), specifically that farmers in an informal risk sharing network who are able to offer each other protection for idiosyncratic risk respond to idiosyncratic basis risk by offering compensation for basis risk. This finding highlights the absence of supplementarity, particularly, where coordinating reciprocal transfers may be difficult or impossible, as is the case when farmers are prescribed into networks.

To verify whether non-indemnity for basis risk is indeed worse for recipients paired with donors from different iddirs, we turn to Table 5.4. Comparing the BPTs (farmers in pre-existing networks) and BPTd (farmers in prescribed networks) sets of models indeed confirms the absence of supplementarity for farmers who are prescribed into networks. For BPTd, the availability of index insurance results in a statistically significant reduction in transfers at a 5% level of significance. In contrast, transfers for BPTs are not significantly different from transfers without formal insurance. Furthermore, when recipients take up index insurance, donors from different iddirs transfer significantly less than when there is no formal insurance. Hence, we find evidence of transfers decreasing rather than increasing in response to basis risk. This suggests that donors punish recipients from different iddirs for taking up index insurance. On the other hand, the coefficients of index rejection are negative and highly significant for both same iddir pairs and different iddir pairs. This suggests that the punishment for rejecting index insurance is as severe for same iddir pairs as for different iddir pairs.

Returning to Table 5.2, when a recipient rejects index insurance, a donor transfers less than in the control. The estimated odds of a transfer are $e^{-1.47458} = 0.23$ times the odds in the control. Hence donors are 77% less likely to make a transfer. The net payoff for the control group and for index rejection is equal at 28 ETB, which means that a recipient's need for a transfer is no less when they reject index insurance compared to when index is unavailable, yet there is a large reduction in transfers. As it is not motivated by a change in the need for a transfer, the significant reduction in transfers may be indicative of a punishment motive for rejecting index insurance.

When there is an aggregate shock

When index insurance is available, transfers are not significantly different from the control group.

When index insurance is taken up, donors transfer less than in the control. The estimated odds of a transfer are $e^{-2.3742} = 0.09$ times the odds of a transfer in the control. This is an expected and warranted decrease as recipients have a net payoff of 82 ETB, which implies a diminished need for a transfer.

When index insurance is rejected, transfers are no different from transfers in the control. This finding is justified by the equality in net payoffs. In contrast to the strong punishment for rejecting index when there is no aggregate shock, donors neither punish nor reward recipients when there is an aggregate shock.

5.3.2 The effect of Indemnity Insurance

The coefficients β_1, in the first row of Table 5.2 measure the effect of indemnity insurance. This coefficient falls away if a donor did not receive the indemnity treatment (Trt2) as shown in (5.11). The estimated effects suggest that the cumulative probabilities, starting at the lower end of the transfers scale increase when indemnity insurance is available, with or without an aggregate shock, decrease with take-up, with or without an aggregate shock, and increase with indemnity rejection when there is an aggregate shock but decrease when there is no aggregate shock.

When there is no aggregate shock

When indemnity insurance is available, transfers are not significantly different from transfers in a network without indemnity insurance.

When indemnity insurance is taken up, donors transfer less than the control. The estimated odds of a transfer are $e^{-2.9283} = 0.05$ times the odds of a transfer in the control. Donors are 95% less likely to make a transfer. Taking up indemnity insurance results in a net payoff of 70 ETB, which diminishes the need for a transfer compared to 28 ETB in the control of no formal insurance. Recall that we defined replacement as a warranted decrease in transfers. A recipient who takes up indemnity insurance and experiences a loss has a reduced need for a transfer. Hence the reduction in transfers when a recipient takes up indemnity insurance is warranted, thus the payout in indemnity insurance replaces transfers in an informal risk sharing network.

When indemnity is rejected, donors transfer less than in the control. The estimated odds of a transfer are $e^{-0.48365} = 0.62$ times the odds of a transfer in the control group. The net payoff of 28 ETB in the control is equal to the net payoff of rejecting indemnity, which means that a recipient's need for a transfer is no less when he or she rejects indemnity compared to when indemnity is unavailable. Therefore, the significant reduction in transfers is indicative of a punishment motive.

When there is an aggregate shock

When indemnity insurance is available transfers are higher than transfers in the control group. The estimated odds of a transfer are $e^{1.1859} = 3.27$ times the odds of a transfer in the control. This finding is unexpected. The net payoff of either 28 ETB or 70 ETB implies no greater need for a recipient to receive a transfer compared to the control. This unexpected finding may be motivated by the fact that there now exists a higher chance of idiosyncratic risk ($\frac{2}{3}$) than in the case of no aggregate shock ($\frac{1}{3}$).

When indemnity insurance is taken up donors transfer less than in the control. The estimated odds of a transfer are $e^{-2.2593} = 0.10$ times the odds of a transfer in the control. This result is understandable because recipients have a net payoff of 70 ETB, which implies a diminished need for a transfer.

When indemnity is rejected transfers are not significantly different from the control, which is justified by the equality in net payoffs. In contrast to the statistically significant evidence of punishment when there is no aggregate shock, donors neither punish nor reward recipients for rejecting indemnity insurance when there is an aggregate shock.

5.3.3 The effect of Pair-Type

The coefficients β^S, in the third row of Table 5.2 measure the effect of Pair-Type. This coefficient is multiplied by s as shown in (5.11). Now, s takes a value 0 when the donor and recipient are from different iddirs and a value of 1 when a donor and recipient are from the same iddir. Therefore the coefficient gives a measure of the extent to which transfers in same iddir pair-types differ from transfers in different iddir pair-types. The coefficients are identical when the baseline category is changed to Trt2 as shown in Table 5.3.

The findings suggest that the cumulative probability starting at the low end of the transfers at 0 ETB is higher for same iddir pair-types than different iddir pair-types although the effect is only statistically significant at 5% when a recipient rejects indemnity insurance and an aggregate shock occurs. In which case, given the iddir, enumerator and treatment, the estimated odds of a transfer in any category to a recipient from the same iddir as the donor is $e^{0.791} = 2.21$ times the odds of a transfer to a recipient from a different iddir.

5.3.4 Checking for robustness

As previously stated, in almost all treatment conditions the predictors Nr.Adults, Education and Occupation proved statistically insignificant at a 5% level of significance. When fitting a CLMM to the subsample with no aggregate shock (B), the addition of Occupation resulted in an unidentified variance-covariance matrix, while the addition of Education resulted in an unidentified variance-covariance matrix when fitting to the subsample with an aggregate shock (Y). However, our model of choice, with only Treatment and Pair-Type effects has a lower AIC than that the corresponding models with Education or Occupation. Hence, the failure to precisely check for robustness with the inclusion of these predictors in either case does not raise much concern.

The effect of insurance availability when there is no aggregate shock

Introducing index insurance results in statistically significant lower transfers compared to the control of informal risk sharing without formal insurance. At a 5% level of significance, the effect of index insurance availability is significantly negative. However, neither effect of indemnity availability nor Pair-Type are statistically significant.

The negative effect of index availability is robust to the addition of Nr.Adults and Education while the coefficient of indemnity insurance availability becomes marginally significant with a p-value of 0.09.

The effect of insurance take-up when there is no aggregate shock

Both index take-up and indemnity take-up have statistically significant negative effects on transfers, while Pair-Type is insignificant. Both of the significant effects are robust to Nr.Adults, Education and Nr.adults-Education interaction.

The effect of insurance rejection when there is no aggregate shock

Both index rejection and indemnity rejection have statistically significant negative effects on transfers, while Pair-Type is statistically insignificant. The effect of index rejection is robust to the addition of Nr.Adults, Education and Nr.adults-Education interaction. Contrarily, the coefficient of indemnity rejection is not robust to the addition of Nr.Adults. The effect becomes only marginally significant with p-value = 0.11.

The effect of insurance availability when there is an aggregate shock

While indemnity availability has a statistically significant positive effect, the effects of both index availability and Pair-Type are insignificant. Furthermore, the coefficient measuring the effect of indemnity availability is robust to the addition of both Nr.adults and Occupation as well as their interaction.

The effect of insurance take-up when there is an aggregate shock

Both indemnity take-up and index take-up have highly statistically significant negative effects, which are robust to Nr.adults, Occupation and Nr.adults-Occupation interaction, whereas Pair-Type is insignificant.

The effect of insurance rejection when there is an aggregate shock

While indemnity rejection and index rejection are statistically insignificant, Pair-Type is statistically significant and robust to to Nr.adults, Occupation and Nr.adults-Occupation interaction.

Chapter 6

Conclusion

Scholarship on index insurance and informal risk-sharing finds a mutually beneficial relationship between index insurance and informal risk sharing when index insurance carries basis risk and provides protection against the aggregate components of a shock while informal risk-sharing covers the idiosyncratic components. This is premised on reciprocity in the networks and is defined as complementarity. This rationale of complementarity is widely (mis)used in agricultural insurance interventions that increasingly offer index insurance on condition of farmers being members of a "group." In such cases farmers are prescribed into groups. Studies have looked at complementarity in pre-existing groups, but the extent to which informal risk-sharing and index insurance are mutually beneficial in prescribed groups is unknown. Therefore this study makes a contribution to this growing literature through examining supplementarity, a more narrowly defined mutually beneficial relationship between index insurance and informal risk sharing in pre-existing and prescribed groups.

The aim of the study was to investigate the effect of formal individual agricultural insurance on transfers by farmers in informal risk sharing networks. Both index and indemnity insurance are investigated for a more holistic picture but the focus is on index insurance. We treat the transfers as categorical data and use multilevel statistical analysis methods. Hence we fit a Cumulative Link Mixed-effects Model (CLMM) using the *clmm* function in the Ordinal package by Rune H. B. Christensen [2014, 2013].

We use a multi-stage sampling design in a framed field experiment which simulated conditions whereby insurance was either available, taken up or rejected. The study compares the levels and odds of transfers in these three conditions to transfers in a control in which formal insurance is unavailable. Secondly, the study compares transfers made by farmers in networks that pre-date the introduction of index insurance (pre-existing) to transfers made by farmers in groups that farmers are prescribed into, in order to access insurance.

Having explored how donor farmers' transfers change in response to a recipient farmer's insurance decision, we then tentatively explore the underlying motives. We focus only on socio-economic incentives rather than preference-dependent factors because economic development interventions alter the former, not the latter. Based on this, we posit that the resultant economic incentives motivate donors to either reward or punish their partners for their insurance decisions.

We find that transfers decrease when index insurance is available, taken up or rejected, with or without an aggregate shock. Farmers are 45% less likely to make a transfer when index insurance is available, 48% less likely when index insurance is taken up and 77% less likely when index insurance is rejected. When there is no aggregate shock and index insurance is available, a recipient's need for a transfer is either heightened or equal to the need in the control group, yet we find a significant reduction in transfers. This suggests that transfers are motivated by more than just the needs of their network members.

When there is no aggregate shock, farmers who take up index insurance suffer basis risk, which exacerbates the need for a transfer. Their resultant net payoff would warrant an increase in transfers relative to the control. Yet, farmers do not respond by supplementing the net payoff through offering compensation for basis risk in index insurance. In fact, this finding of an unwarranted decrease in transfers suggests a punishment motive underlying the transfer. This motive may not be unrealistic as donor farmers may be punishing recipient farmers for taking on additional idiosyncratic risk. We further find that this punishment or non-compensation for basis risk is worse for farmers in prescribed groups than farmers in pre-existing groups.

When there is no aggregate shock, the net payoff for the control group and for farmers who had rejected index insurance is equal at 28 Ethiopian Birr (ETB), which means that a recipient's need for a transfer is no less when he or she rejects index insurance compared to when index is unavailable. As it is not motivated by a change in the need for a transfer, the significant reduction in transfers may be indicative of punishment for rejecting index insurance. We also find that the punishment for rejecting index insurance is equally severe for same iddir pairs and different iddir pairs. In contrast to the strong punishment for rejecting index insurance when there is no aggregate shock, donors neither punish nor reward recipients for rejecting index insurance when there is an aggregate shock.

Turning to indemnity insurance, we find that when there is no aggregate shock donors are 95% less likely to make a transfer when indemnity insurance is taken up. Taking up indemnity insurance results in a net payoff of 70 ETB, which diminishes the need for a transfer compared to 28 ETB in the control of no formal insurance. Thus the payout from indemnity replaces transfers in informal risk sharing networks. Meanwhile, donors are 38% less likely to make a transfer when indemnity is rejected, relative to the control group. The net payoff of 28 ETB in the control is equal to the net payoff of rejecting indemnity, which means that the recipient's need for a transfer is no less. Therefore, the significant reduction in transfers suggests an underlying punishment motive. In contrast, when there is an aggregate shock, there is no evidence of donors punishing recipients for rejecting indemnity insurance.

Finally, we find that in terms of both the levels transferred and the odds of a transfer, farmers in pre-existing groups are better off than farmers in prescribed groups, particularly when an an aggregate shock occurs and a recipient farmer has rejected insurance.

The choice to focus only on socio economic dependent factors when linking the transfers to their likely underlying motives is a weakness of this study. Preference-dependent factors offer alternative transfer motives. In order to design agricultural insurance intervention programs that improve the financial resilience of low-income farmers it its critical to understand how the complex interaction of both preference-dependent factors and socio-economic-dependent factors affect the extent to which formal insurance and informal risk sharing networks can be mutually beneficial. This area presents a fertile area of research.

This study makes a novel contribution to scholarship on the nexus between formality and informality in smallholder farmers' risk management strategies through its empirical findings. It adds to our understanding of how agricultural insurance interventions impact transfer behaviour in informal risk-sharing networks by way of altering socio-economic incentives. Providers of micro insurance in low-income farming markets will find the empirical findings useful in that they show the need for care when designing agricultural insurance interventions aimed at realising a mutually beneficial relationship between index insurance and informal risk sharing. The findings also highlight the need for caution in prescribing farmers into groups, particularly when an intervention aims to leverage off of mechanisms at play in the informal networks.

Appendices

Appendix A

List of variables

Variable name	Variable description	Values NA = missing
Rnr	Respondent number on PDA and on paper interviews	
Rnrnew	Corrected respondent number	
Trt	Treatment number	T1 T2 T3 T4
Year	Year in which the experiments was done	2014
Month	Month in which the experiment was done	February = 2 March = 3
Day	Day of the month in which experiment was done	
Sessionlead	Session leaders, one of three	Karlijn = K Qhelile = Q Tagel = T
Enum	Name of the Enumerator	Abr Ashu Beky Gech Gir Haf Kib Mul B Mul K Samy NA = missing
Session	Session number - 1 to 48	
Idd_nr	Iddir number	Factor with 16 levels idd1 - idd16
THE QUESTIONS NUMBERED ET ARE EDUCATION TESTING, TO MEASURE THE FARMERS' LEVEL OF UNDERSTANDING OF THE GAME		
ET2.1	If farmer rolls a yellow dice (bad crop) how much is the value of his crops in birr? (right answer = 28)	Integers
ET2.2 If crop is bad (yellow dice) but weather is good (blue token)→ How much does farmer think is …….with crop insurance:		
ET2.2_1	His/her insurance claim? O 0 birr O 72 birr (right answer=72)	Integers
ET2.2_2	His/her eventual payoff?…………BIRR (right answer= 28 – 30 = 70)	Integers
ET2.3 If crop is good (blue dice) and weather is good (blue token)→ How much does farmer think is …….with crop insurance		
ET2.3_1	His/her insurance claim? O 0 birr O 72 birr (right answer=0)	Integers
ET2.3_2	His/her eventual payoff? …………BIRR (right answer= 100-30= 70)	Integers
ET2.4 If crop is bad (yellow dice) and weather is bad (yellow token)→ How much does farmer think is …….with crop insurance:		
ET2.4_1	His/her insurance claim? O 0 birr O 72 birr	Integers

	(right answer=72)	
ET2.4_2	His/her eventual payoff? …………BIRR (right answer= 28-30+72= 70)	Integers
ET2.5	How much is the insurance premium? ……BIRR (right answer = 30)	Integers
ET3.1	If farmer rolls a yellow dice (bad crop) how much is the value of his crops in birr? …………..BIRR (right answer = 28)	Integers
ET3.2 If crop is bad (yellow dice) but weather is good (blue token)→ How much does farmer think is …….with rainfall insurance:		
ET3.2_1	His/her insurance claim? O 0 birr O 72 birr (right answer=0)	Integers
ET3.2_2	His/her eventual payoff?…………BIRR (right answer= 28 – 18 = 10)	Integers
ET3.3 If weather is bad (yellow token) and crop is good (blue dice) → How much does farmer think is …….with rainfall insurance:		
ET3.3_1	His/her insurance claim O 0 birr O 72 birr (right answer= 72)	Integers
ET3.3_2	His/her eventual payoff ………..BIRR (right answer = 100 – 18 + 72 = 154)	Integers
ET3.4 If weather is good and crop is good → How much does farmer think is …….with rainfall insurance:		
ET3.4_1	His/her insurance claim O 0 birr O 72 birr (right answer= 0)	Integers
ET3.4_2	His/her eventual payoff ………..BIRR (right answer = 100 – 18 = 82)	Integers
ET3.5 If weather is bad and crop is bad → How much does farmer think is …….with rainfall insurance:		
ET3.5_1	His/her insurance claim O 0 birr O 72 birr (right answer= 72)	Integers
ET3.5_2	His/her eventual payoff ………..BIRR (right answer = 28-18 + 72 = 82)	Integers
ET3.6	How much is the insurance premium? ………Birr (right answer = 18 Birr)	Integers
ET4.1	ET4.1 If farmer rolls a yellow dice (bad crop) how much is the value of his crops in birr? (right answer = 28)	Integers
ET4.2 If crop is bad (yellow dice) but weather is good (blue token) How much does farmer think is …….with rainfall insurance:		
ET4.2_1	- His/her insurance claim? O 0 birr O 72 birr (right answer=0)	Integers
ET4.2_2	- His/her eventual payoff?…………BIRR (right answer= 28 – 18 = 10)	Integers
ET4.3 If weather is bad (yellow token) and crop is good (blue dice) → How much does farmer think is …….with crop insurance:		
ET4.3_1	His/her insurance claim O 0 birr O 72 birr (right answer=0)	Integers
ET4.3_2	His/her eventual payoff ………..BIRR (right answer = 100 – 30 = 70)	Integers

ET4.4 If weather is bad (yellow token) and crop is good (blue dice) → How much does farmer think is …….with rainfall insurance:		
ET4.4_1	His/her insurance claim O 0 birr O 72 birr (right answer= 72)	Integers
ET4.4_2	His/her eventual payoff ………..BIRR (right answer = 100 – 18 + 72 = 154)	Integers
ET4.5	How much is the rainfall insurance premium?........BIRR (right answer = 18 Birr	Integers
ET4.6	How much is the crop insurance premium?........BIRR (right answer = 30 Birr)	Integers
I_or_J	Is this farmer i or farmer j?	I J
PT	Pair-Type: Was the farmer paired with someone from the Same iddir or a Different iddir?	S = same iddir D = different iddir
AS	Aggregate Shock: What token was drawn to determine the weather?	B = blue token Y = yellow token
Tij_GEN	How much do you want to transfer to j? *Asked in all treatments and depends on expectation of j's behaviour*	Integers
T1ij.3_S	How much do you want to transfer to j in case j is from the same iddir? *Only asked in T1*	Integers
T1ij.3_D	How much do you want to transfer to j in case j is from another iddir? *Only asked in T1*	Integers
T4ij_No_Ins	How much do you want to transfer to j in case j did not get an insurance offer? *Only asked in T4*	Integers
Tij_C_Y	How much do you want to transfer to j in case j received crop-insurance offer and took crop insurance? *Asked in T2 and T4*	Integers
Tij_C_No	How much do you want to transfer to j in case j received crop-insurance offer and DID NOT take crop insurance? *Asked in T2 and T4*	Integers
Tij_R_Y	How much do you want to transfer to j in case j received rain insurance offer and took rain insurance? *Asked in T3 and T4*	Integers
Tij_R_No	How much do you want to transfer to j in case j received rain-insurance offer and DID NOT take rain insurance? *Asked in T3 and T4*	Integers
LiKe_C_Ins	How likely do you think it is that farmer j will have taken crop insurance? *Asked in T2 and T4*	1 - 10
LiKe_R_ins	How likely do you think it is that farmer j will have	1 - 10

	take rainfall insurance *Asked in T3 and T4*	
RnrPDA_new	New survey identifier from PDA	
EnumPDA	Enumerator PDA	Abr Ashu Beky Gech Gir Haf Kib Mul B Mul K Samy NA = missing
SessionPDA	Session number	1-48
Tabia	Local administrative unit	NA = missing
Name	Name of farmer	Name
Hhd_head	Is the farmer the head of the household?	Yes No NA = missing
Gender	Gender of farmer	Woman Man NA = missing
Age	Age of farmer	NA = missing
Marital	Household head marital status	Divorced Married Single Separated Widowed Other NA = missing
Lit	Is the household head literate?	Yes No NA = Missing
Edu	What is the level of education of the household head?	Religious education Primary incomplete Primary complete Junior complete High School incomplete High School complete No education Adult education NA = missing
Occ	What is the occupation of the household head?	Farmer Laborer Trader Laborer; Farmer Trader; Farmer All of above NA = missing
Nr_Adult	Nr of adults in the household	An outlier at 32

		NA = missing
Nr_Child	Nr of children in the household	NA = missing
Livestock	Does the farmer have livestock?	Yes No NA = missing
Cows	Nr of	NA = missing
Bulls	Nr of	NA = missing
Oxen	Nr of	NA = missing
Sheep	Nr of	NA = missing
Goats	Nr of	NA = missing
Beehives	Nr of	Outlier at 100 NA = missing
Donkeys	Nr of	NA = missing
Horses	Nr of	NA = missing
Mules	Nr of	NA = missing
Camels	Nr of	NA = missing
Chickens	Nr of	NA = missing
Irrigate	Does farmer have irrigated land	No Yes, without water pump Yes, with water pump Yes NA = missing
Idd_num	How many iddirs is the farmer a member of?	0 or 1 Outlier at 11 Outlier at 50 NA = missing
Idd_name	Name of iddir	NA = missing
Idd_yrs	Number of years a member of the iddir?	1 - 98 Outlier at 120 NA = missing
Idd_contrib	How often does the farmer contribute in the iddir?	At start Monthly Shock = If there is a shock At start; Monthly Shock, Monthly All above cases NA = missing
Idd_monthly	How much does the farmer contribute monthly in the iddir in ETB (local currency)	
Idd_shock	How much does the farmer contribute if there is a shock (ETB)	
Idd_support	How many times has the farmer received support from the iddir in the last 5 years?	
Farm_own	Does the farmer own farm land?	Yes No NA = missing
Tsem	If so, how many tsemidi (0.25ha)	
Farm_irr	Does farmer have any access to irrigation?	Yes, with a pump Yes, but with no pump No NA = missing

Tsem_irr	How many Tsemdi is irrigated?	
Loss50_idd	On average how many years in the last ten years did the farmers in your iddir experience more that 50% crop loss?	1-10
Loss50_own	How many years in the last ten years on average did you experience more that 50% crop loss?	1-10
Loss25_idd	How many years in the last ten years did the farmers in your iddir experience between 25% to 50% crop loss?	1-10
Loss25_own	How many years in the last ten years did you experience between 25% to 50% crop loss?	1-10

Appendix B

Household survey

ET-Final2 Questionnaire

Farmer number ………………………

HOUSEHOLD INFORMATION

1. Which Tabia are you from? ……………………………………………..........

2. What is your name…………………………………………………………

3. Are you the head of your household? ……………O Yes O No

4. **Sex** ……………………………….. O Man O Woman

5. Age ……….

6. What is the marital status of the household head?
O Divorced O Married O Single O Separated O Widowed O Other

7. Can the household head read and write a simple message in any language?.........O Yes O No

8. What is the level of education of your household head?
O Primary incomplete O Adult education O Primary complete O Junior complete
O Religious education O High School O High School Complete O No education

9. What is the major occupation of the household head? (MARK ALL THAT APPLY)
O Trader O Labourer O Farmer

10. How many adults are currenly living in your household?......................................

11. How many children are currently living in your household?..................................

WEALTH INFORMATION

1. Do you have livestock? ……………O Yes O No

2. How many of each does the farmer own?

- Cows ……………………..
- Bulls ………………………
- Oxen ……………………..
- Sheep ……………………
- Goats …………………….
- Beehives ………………..
- Donkeys ………………...
- Horses …………………..
- Mules ……………………
- Camels …………………
- Chickens ………………..

3. Do you have access to irrigation? ……………O Yes O No

IDDIR INFORMATION

1. How many iddirs are you a member of?..

2. Which iddirs are you a member of?
i) ……………………………………………………………………
ii) ……………………………………………………………………
iii) ……………………………………………………………………..

3. How many years have you been a member of your iddir?.......................

4. How often do you contribute to the iddir? (TICK ALL THAT APPLY)
O At start O If there is a shock O Monthly

5. How much money do contribute to the iddir monthly (Birr)? ……………….

6. How much money do you contribute when a member in your iddir has a shock (Birr)?

7. How many times have received monetary or in-kind support from your iddir in the last 5 years?.....................

FARMING INFORMATION

1. Are you the main person responsible for farming? ……………O Yes O No

2. If so, how many tsemidi (0.25ha) do you own?.....................
→ OR How many tsemidi do you rent?.......................

3. Do you have any access to irrigation?
O Yes, with a pump ……………O Yes, but with no pump O No

4. How many tsemidi is irrigated? ………………..

5. Instruct the farmer to place a number of tokens. Explain to the farmer that he/she responds to the questions in this section by placing a number of tokens out of 10 onto the table, where 0 means not likely or possible, 5 means somewhat likely and 10 means certain

6. On average how many years in the last ten years did the farmers in your iddir experience more that 50% crop loss? …………………………..

7. How many years in the last ten years on average did you experience more that 50% crop loss? …………………………

8. How many years in the last ten years did the farmers in your iddir experience between 25% to 50% crop loss? ……………………………

9. How many years in the last ten years did you experience between 25% to 50% crop loss? ………………………………..

RISK PERCEPTION QUESTIONS

1. Explain to the farmer that he/she responds to the questions in this section by placing a number of tokens out of the 10 they have onto the table –

R1. How willing are you to take risks on a scale from 0 (avoid risks) to 10 (always take risks)……..

R2. How willing are you to take risks on a scale from 0 (avoid risks) to 10 (always take risks) when deciding about loans?...............

R3. How willing are you to take risks on a scale from 0 (avoid risks) to 10 (always take risks) when buying seedlings?..................

R4. How willing are you to take risks on a scale from 0 (avoid risks) to 10 (always take risks) when buying fertilizer ?....................

R5. How willing are you to take risks when negotiating crop prices?....................

R6. How would you rate your willingness to take risks when negotiating over share cropping with other people? ………………

R7. How likely are you to plant terf or wheat (risky crops) in the next season?.........................

R8. How likely are you to spend a large amount of your income in the brewery?..……………......

R9. How risky do you think it is to plant terf or wheat (risky crops) in the next season?..............

R10. How risky do you think it is to spend a large amount of your income in the brewery?..........

R11. How risky do think it is to bet a large portion fof your income on endomino?.......................

Appendix C

The risk-sharing games

Risk-sharing questionnaire T1

Date:

Session:

Session leader

Name enumerator:

Farmer number:

Before starting check (tick box on all):	
O	1 questionnaire
O	1 red dice; 1 white dice in cup
O	1 white envelope → empty
O	1 brown envelope → 200 birr
O	PDA
O	'Risk-sharing sheets'
O	'Risk-preference sheets'
O	Pencil
O	Winnings cards
O	10 coins for risk perception questions

ASSIGN NUMBER TO FARMER → WRITE ON WINNINGS CARD, SMALL CARD AND QUESTIONNAIRE

ASSIGN ENUMERATOR TO FARMER

EXPLAIN WE PLAY TWO GAMES IN WHICH THEY CAN EARN MONEY. PAYOFF AT THE END WILL BE A DRAW FROM THESE TWO GAMES. MINIMUM 50 BIRR; MAXIMUM 160 BIRR

ENUMERATOR HAS BROWN ENVELOPE WITH 200 BIRR

EDUCATION: EXPERIENCING WEATHER AND CROP LOSS

TWO PRACTICE ROUNDS. FARMERS WHO GET BLUE GET 100 BIRR; YELLOW 28 BIRR.

EDUCATION RISK-SHARING

PLAY RISK-SHARING WITH ENUMERATORS AS EXAMPLE.

FARMERS WILL BE TEAMED UP (LIKE THE ENUMERATORS) WITH SOMEONE IN THE ROOM.

AT END RECOLLECT 200 BIRR FROM FARMERS

T1; farmer i and j

FARMER HAS EMPTY ENVELOPE. TELL FARMER WINNINGS WILL BE COLLECTED IN ENVELOPE.

NOW WE ARE GOING TO PLAY THE REAL GAME

TELL FARMER IF HE/SHE IS FARMER i or FARMER j O i O j

TELL FARMER IF HE/SHE IS TEAMED-up with farmer from ….. O same iddir O different iddir

SHOW LIST OF WHO THEY ARE TEAMED UP WITH.

T1; farmer i

T1i.1 Weather (token draw) O drought (yellow) O medium/good weather (blue)

PAY FARMER i 100 BIRR FROM BROWN ENVELOPE

REMIND FARMER THAT J MAY EXPERIENCE A LOSS.

[REMEMBER T1i.2 = T1i.3 or T1i.4]

T1i.2 How much does i want to share with j in case j has bad crop and receives only 28 birr?

…………birr [**RECORD ON WINNINGSCARD]**

T1i.3 How much does i want to share with j in case j has bad crop and receives only 28 birr and is from the same iddir?

…………birr [**RECORD ON WINNINGSCARD]**

T1i.4 How much does i want to share with j in case j has bad crop and receives only 28 birr and is not from the same iddir?

…………birr [**RECORD ON WINNINGSCARD]**

[BRING WINNINGSCARD TO SESSION LEADER]

T1; farmer j

T1j.1 Weather (token draw) O drought (yellow) O medium/good weather (blue)

T1j.2 Farmer dice roll O 100 BIRR (good crop) O 28 BIRR (bad crop)

PAY THE FARMER 100 or 28 BIRR FOR THEIR CROP

[RECORD ON WINNINGSCARD]

[BRING WINNINGSCARD TO SESSION LEADER]

TELL FARMERS WE WILL CALCULATE:

- **FOR i: THEIR PAYOFF IN THE GAME BASED ON FARMER j OUTCOME (100 – POTENTIAL TRANSFER TO j IF J HAS LOSS**
- **FOR j: THEIR PAYOFF BASED ON THEIR DICE ROLL (WHICH THEY HAVE BEEN GIVEN) + FARMER i TRANSFER (DETERMINED LATER)**

RISK PREFERENCES GAME 2:

*******BEFORE EVERY GAME: REMIND FARMER THAT PAYOFF AT END DEPENDS ON OUTCOMES IN THIS GAME!!!!!********

G1.1	#	L	S	Sure	G1.2	#	L	S	Sure	G1.4	#	L	S	Sure
Lottery	1	O	O	0	Lottery	1	O	O	0	Lottery	1	O	O	0
	2	O	O	10		2	O	O	10		2	O	O	10
Win 160,	3	O	O	20	Win 160,	3	O	O	20	Win 160,	3	O	O	20
p=0.5	4	O	O	30	p=0.25	4	O	O	30	p=0.75	4	O	O	30
	5	O	O	40		5	O	O	40		5	O	O	40
**********	6	O	O	50	*******	6	O	O	50	***	6	O	O	50
REMIND	7	O	O	60	**REMIND**	7	O	O	60	**REMIND**	7	O	O	60
FARMER	8	O	O	70	**FARMER**	8	O	O	70	**FARMER**	8	O	O	70
THAT	9	O	O	80	**THAT**	9	O	O	80	**THAT**	9	O	O	80
PAYOFF AT	10	O	O	90	**PAYOFF AT**	10	O	O	90	**PAYOFF AT**	10	O	O	90
END	11	O	O	100	**END**	11	O	O	100	**END**	11	O	O	100
DEPENDS	12	O	O	110	**DEPENDS**	12	O	O	110	**DEPENDS**	12	O	O	110
ON	13	O	O	120	**ON**	13	O	O	120	**ON**	13	O	O	120
OUTCOMES	14	O	O	130	**OUTCOMES**	14	O	O	130	**OUTCOMES**	14	O	O	130
IN THIS	15	O	O	140	**IN THIS**	15	O	O	140	**IN THIS**	15	O	O	140
GAME!!!!!	16	O	O	150	**GAME!!!!!**	16	O	O	150	**GAME!!!!!**	16	O	O	150
	17	O	O	160		17	O	O	160		17	O	O	160

*******BEFORE EVERY GAME: REMIND FARMER THAT PAYOFF AT END DEPENDS ON OUTCOMES IN THIS GAME!!!!!********

G1.6	#	L	S	Sure	G1.7	#	L	S	Sure	G1.9	#	L	S	Sure
Lottery	1	O	O	-160	Lottery	1	O	O	-160	Lottery	1	O	O	-160
	2	O	O	-150		2	O	O	-150		2	O	O	-150
Lose 160,	3	O	O	-140	Lose 160,	3	O	O	-140	Lose 160,	3	O	O	-140
P=0.5	4	O	O	-130	p=0.25	4	O	O	-130	P=0.75	4	O	O	-130
	5	O	O	-120		5	O	O	-120		5	O	O	-120
****	6	O	O	-110	********	6	O	O	-110	*******	6	O	O	-110
REMIND	7	O	O	-100	**REMIND**	7	O	O	-100	**REMIND**	7	O	O	-100
FARMER	8	O	O	-90	**FARMER**	8	O	O	-90	**FARMER**	8	O	O	-90
THAT	9	O	O	-80	**THAT**	9	O	O	-80	**THAT**	9	O	O	-80
PAYOFF AT	10	O	O	-70	**PAYOFF AT**	10	O	O	-70	**PAYOFF AT**	10	O	O	-70
END	11	O	O	-60	**END**	11	O	O	-60	**END**	11	O	O	-60
DEPENDS	12	O	O	-50	**DEPENDS**	12	O	O	-50	**DEPENDS**	12	O	O	-50
ON	13	O	O	-40	**ON**	13	O	O	-40	**ON**	13	O	O	-40
OUTCOMES	14	O	O	-30	**OUTCOMES**	14	O	O	-30	**OUTCOMES**	14	O	O	-30
IN THIS	15	O	O	-20	**IN THIS**	15	O	O	-20	**IN THIS**	15	O	O	-20
GAME!!!!!	16	O	O	-10	**GAME!!!!!**	16	O	O	-10	**GAME!!!!!**	16	O	O	-10
	17	O	O	0		17	O	O	0		17	O	O	0

G1.11	blue	Yellow	#	L	S	**Sure**
Mixed Lottery:	160	-160	1	O	O	**0**
Win 160 if blue,	160	-150	2	O	O	
or lose an increasing	160	-140	3	O	O	
amount if yellow	160	-130	4	O	O	
with p=0.5	160	-120	5	O	O	
	160	-110	6	O	O	
	160	-100	7	O	O	
********REMIND FARMER**	160	-90	8	O	O	
THAT PAYOFF AT END	160	-80	9	O	O	
DEPENDS ON OUTCOMES	160	-70	10	O	O	
IN THIS GAME!!!!!	160	-60	11	O	O	
	160	-50	12	O	O	
	160	-40	13	O	O	
	160	-30	14	O	O	
	160	-20	15	O	O	
	160	-10	16	O	O	
	160	0	17	O	O	

Binswanger O A O B O C O D O E O F

3. DETERMINING PAYOFF GAME 2:

3.1 Random draw of one of the 11 price lists (G1.1-G1.11)
O 1 O 2 O 3 O 4 O 5 O 6 O 7 O 8 O 9 O 10 O 11
3.2 Random draw of one of the 17 decision rows decision row: .....
3.3 Look up farmer's choice in drawn price list/row O lottery O sure
if subject chose lottery:
3.4 put tokens price list and draw token: O Blue O Yellow

COLLECT WHITE ENVELOPES

FILL OUT WINNINGSCARD AND GIVE TO SESSION LEADER

NOW START WORKING ON PDA

Before starting check (tick box on all):
O 1 questionnaire
O 1 red dice; 1 white dice in cup
O 1 white envelope → empty
O 1 brown envelope → 200 birr
O PDA
O 'Risk-sharing sheets'
O 'Risk-preference sheets'
O Pencil
O Winnings cards
O 10 coins for risk perception questions

Risk-sharing questionnaire T2

Date:

Session:

Session leader

Name enumerator:

Farmer number:

ASSIGN NUMBER TO FARMER → WRITE ON WINNINGSCARD, SMALL CARD AND QUESTIONNAIRE

ASSIGN ENUMERATOR TO FARMER

EXPLAIN WE PLAY TWO GAMES IN WHICH THEY CAN EARN MONEY. PAYOFF AT THE END WILL BE A

DRAW FROM THESE TWO GAMES. MINIMUM 50 BIRR; MAXIMUM 160 BIRR

ENUMERATOR HAS BROWN ENVELOPE WITH 200 BIRR

EDUCATION: EXPERIENCING WEATHER AND CROP LOSS

TWO PRACTICE ROUNDS. FARMERS WHO GET BLUE GET 100 BIRR; YELLOW 28 BIRR.

EDUCATION RISK-SHARING

PLAY RISK-SHARING WITH ENUMERATORS AS EXAMPLE.

FARMERS WILL BE TEAMED UP (LIKE THE ENUMERATORS) WITH SOMEONE IN THE ROOM.

AT END RECOLLECT 200 BIRR FROM FARMERS

T2; crop insurance; farmer i and j

FARMER HAS EMPTY ENVELOPPE. TELL FARMER WINNINGS WILL BE COLLECTED IN ENVELOPPE.
MAKE SURE YOU HAVE ENVELOPE WITH 200 BIRR
EDUCATION: EXPERIENCING CROP INSURANCE
PLAY RISK-SHARING WITH CROP INSURANCE WITH ENUMERATORS AS EXAMPLE

ASK UNDERSTANDING QUESTIONS:

ET2.1 If farmer rolls a yellow dice (bad crop) how much is the value of his crops in birr? (right answer = 28)

ET2.2 If crop is bad (yellow dice) but weather is good (blue token)→ How much does farmer think is …….with crop insurance:

- His/her insurance claim?	O 0 birr	O 72 birr	(right answer=72)
- His/her eventual payoff?	…………BIRR		(right answer= 28 – 30 = 70)

ET2.3 If crop is good (blue dice) and weather is good (blue token)→ How much does farmer think is …….with crop insurance:

- His/her insurance claim?	O 0 birr	O 72 birr	(right answer=0)
- His/her eventual payoff?	…………BIRR		(right answer= 100-30= 70)

ET2.4 If crop is bad (yellow dice) and weather is bad (yellow token)→ How much does farmer think is …….with crop insurance:

- His/her insurance claim?	O 0 birr	O 72 birr	(right answer=72)

- His/her eventual payoff? ………..BIRR (right answer= 28-30+72= 70)
ET2.5 How much is the insurance premium? ……BIRR (right answer = 30)

PLEASE CHECK IF YOU HAVE 200 BIRR IN ENVELOPE.

WE ARE NOW GOING TO PLAY THE REAL GAME

TELL FARMER IF HE/SHE IS FARMER i or FARMER j O i O j

GIVE FARMER j 30 BIRR TO PUT IN WHITE ENVELOPE
TELL FARMER IF HE/SHE IS TEAMED-up with farmer from ….. O same iddir O different iddir

SHOW LIST OF WHO THEY ARE TEAMED UP WITH.

T2; farmer j; TREATMENT CROP INSURANCE AND RISK-SHARING

TELL FARMER THAT IF HE/SHE WANTS TO TAKE INSURANCE THE PREMIUM (30 BIRR) WILL BE DEDUCTED FROM THEIR PAYOFF AT THE END. (100-30 OR 28-30+72 + TRANSFER) **[SHOW ON RISK-SHARING SHEETS]**

T2j.1Does farmer want to take insurance? O yes O no **[RECORD ON WINNINGSCARD]**

LET THE FARMER PAY THE PREMIUM OF 30 BIRR IN CASE YES.

T2j.2 Weather (token draw) O drought (yellow) O medium/good weather (blue) **[RECORD ON WINNINGSCARD]**

PAY FARMER i 100 BIRR

T1j.3 Farmer dice roll O 100 BIRR (good crop) O 28 BIRR (bad crop)

PAY THE FARMER 100 or 28 BIRR FOR THEIR CROP

PAY THE FARMER 72 BIRR in case dice roll is 28 BIRR

[RECORD ON WINNINGSCARD]

[BRING WINNINGSCARD TO SESSION LEADER]

T2; farmer i; TREATMENT CROP INSURANCE AND RISK-SHARING

(J WILL DECIDE IF HE/SHE WANTS TO INSURE FIRST)

T2i.1 Weather (token draw) O drought (yellow) O medium/good weather (blue)

PAY FARMER i 100 BIRR FROM BROWN ENVELOPPE

REMIND FARMER THAT J MAY EXPERIENCE A LOSS.

T2i.2 **(READ NOTE TO ENUMERATOR FIRST)** How much does i want to share with j in case j has bad crop [FARMER i KNOWS THAT FARMER J WAS OFFERED INSURANCE BUT DOESN'T KNOW IF HE/SHE TOOK IT]? ……………BIRR **[RECORD ON WINNINGSCARD]**

NOTE TO ENUMERATOR: SHOW ON RISK-SHARING SHEET (BASED ON TOKEN DRAWN) THE TWO POSSIBILITIES (j has 70 birr if taken insurance or j has 28 birr if not taken insurance)

T2i.3 **(READ NOTE TO ENUMERATOR FIRST)** How much does i want to share with j in case j has bad crop and took insurance?...........BIRR

NOTE TO ENUMERATOR: SHOW ON RISK-SHARING SHEET (BASED ON TOKEN DRAWN) THE PAYOFF OF j (if j took insurance and has shock j will have 70 birr)

T2i.3A What best describes this choice? O compensate because bad crop O compensate less because insurance O punish because take insurance O reward because take insurance

T2i.4 **(READ NOTE TO ENUMERATOR FIRST)** How much does i want to share with j in case j has bad crop and took no insurance?...........BIRR

NOTE TO ENUMERATOR: SHOW ON RISK-SHARING SHEET (BASED ON TOKEN DRAWN) THE PAYOFF OF j (if j did not take insurance and has shock j will have 28 birr)

T2i.4A What best describes this choice? O compensate because bad crop O punish because not take insurance

T2i.5 How likely, on a scale from 0-10 (10 likely, 0 unlikely) does farmer i think it is that farmer j will take insurance?

O 0 O1 O2 O3 O4 O5 O6 O7 O8 O9 O10

TELL FARMERS WE WILL CALCULATE:

- **FOR i: THEIR PAYOFF IN THE GAME (100 – TRANSFER) BASED ON FARMER j'S**
    - **Insurance decision (yes or no)**
    - **Dice roll (yellow or blue)**
- **FOR j: THEIR PAYOFF BASED ON FARMER i TRANSFER FOR EITHER T2i.2 OR T2i.3 OR T2i.4**

RISK PREFERENCES GAME 2:

*******BEFORE EVERY GAME: REMIND FARMER THAT PAYOFF AT END DEPENDS ON OUTCOMES IN THIS GAME!!!!!********

G1.1	#	L	S	Sure	G1.2	#	L	S	Sure	G1.4	#	L	S	Sure
Lottery Win 160, p=0.5 ********** **REMIND FARMER THAT PAYOFF AT END DEPENDS ON OUTCOMES IN THIS GAME!!!!!**	1	O	O	0	Lottery Win 160, p=0.25 ******* **REMIND FARMER THAT PAYOFF AT END DEPENDS ON OUTCOMES IN THIS GAME!!!!!**	1	O	O	0	Lottery Win 160, p=0.75 *** **REMIND FARMER THAT PAYOFF AT END DEPENDS ON OUTCOMES IN THIS GAME!!!!!**	1	O	O	0
	2	O	O	10		2	O	O	10		2	O	O	10
	3	O	O	20		3	O	O	20		3	O	O	20
	4	O	O	30		4	O	O	30		4	O	O	30
	5	O	O	40		5	O	O	40		5	O	O	40
	6	O	O	50		6	O	O	50		6	O	O	50
	7	O	O	60		7	O	O	60		7	O	O	60
	8	O	O	70		8	O	O	70		8	O	O	70
	9	O	O	80		9	O	O	80		9	O	O	80
	10	O	O	90		10	O	O	90		10	O	O	90
	11	O	O	100		11	O	O	100		11	O	O	100
	12	O	O	110		12	O	O	110		12	O	O	110
	13	O	O	120		13	O	O	120		13	O	O	120
	14	O	O	130		14	O	O	130		14	O	O	130
	15	O	O	140		15	O	O	140		15	O	O	140
	16	O	O	150		16	O	O	150		16	O	O	150
	17	O	O	160		17	O	O	160		17	O	O	160

*******BEFORE EVERY GAME: REMIND FARMER THAT PAYOFF AT END DEPENDS ON OUTCOMES IN THIS GAME!!!!!********

G1.6	#	L	S	Sure	G1.7	#	L	S	Sure	G1.9	#	L	S	Sure
Lottery Lose 160, P=0.5 **** **REMIND FARMER THAT PAYOFF AT END DEPENDS ON OUTCOMES IN THIS GAME!!!!!**	1	O	O	-160	Lottery Lose 160, p=0.25 ******** **REMIND FARMER THAT PAYOFF AT END DEPENDS ON OUTCOMES IN THIS GAME!!!!!**	1	O	O	-160	Lottery Lose 160, P=0.75 ******* **REMIND FARMER THAT PAYOFF AT END DEPENDS ON OUTCOMES IN THIS GAME!!!!!**	1	O	O	-160
	2	O	O	-150		2	O	O	-150		2	O	O	-150
	3	O	O	-140		3	O	O	-140		3	O	O	-140
	4	O	O	-130		4	O	O	-130		4	O	O	-130
	5	O	O	-120		5	O	O	-120		5	O	O	-120
	6	O	O	-110		6	O	O	-110		6	O	O	-110
	7	O	O	-100		7	O	O	-100		7	O	O	-100
	8	O	O	-90		8	O	O	-90		8	O	O	-90
	9	O	O	-80		9	O	O	-80		9	O	O	-80
	10	O	O	-70		10	O	O	-70		10	O	O	-70
	11	O	O	-60		11	O	O	-60		11	O	O	-60
	12	O	O	-50		12	O	O	-50		12	O	O	-50
	13	O	O	-40		13	O	O	-40		13	O	O	-40
	14	O	O	-30		14	O	O	-30		14	O	O	-30
	15	O	O	-20		15	O	O	-20		15	O	O	-20
	16	O	O	-10		16	O	O	-10		16	O	O	-10
	17	O	O	0		17	O	O	0		17	O	O	0

G1.11	blue	Yellow	#	L	S	**Sure**
Mixed Lottery:	160	-160	1	O	O	**0**
Win 160 if blue,	160	-150	2	O	O	
or lose an increasing	160	-140	3	O	O	
amount if yellow	160	-130	4	O	O	
with p=0.5	160	-120	5	O	O	
	160	-110	6	O	O	
	160	-100	7	O	O	
********REMIND FARMER**	160	-90	8	O	O	
THAT PAYOFF AT END	160	-80	9	O	O	
DEPENDS ON OUTCOMES	160	-70	10	O	O	
IN THIS GAME!!!!!	160	-60	11	O	O	
	160	-50	12	O	O	
	160	-40	13	O	O	
	160	-30	14	O	O	
	160	-20	15	O	O	
	160	-10	16	O	O	
	160	0	17	O	O	

Binswanger O A O B O C O D O E O F

3. DETERMINING PAYOFF GAME 2:

3.1 Random draw of one of the 11 price lists (G1.1-G1.11)
O 1 O 2 O 3 O 4 O 5 O 6 O 7 O 8 O 9 O 10 O 11
3.2 Random draw of one of the 17 decision rows decision row: .....
3.3 Look up farmer's choice in drawn price list/row O lottery O sure
if subject chose lottery:
3.4 put tokens price list and draw token: O Blue O Yellow

COLLECT WHITE ENVELOPES

FILL OUT WINNINGSCARD AND GIVE TO SESSION LEADER

NOW START WORKING ON PDA

Risk-sharing questionnaire T3

Date:

Session:

Session leader

Name enumerator:

Farmer number:

Before starting check (tick box on all):

O 1 questionnaire
O 1 red dice; 1 white dice in cup
O 1 white envelope → empty
O 1 brown envelope → 200 birr
O PDA
O 'Risk-sharing sheets'
O 'Risk-preference sheets'
O Pencil
O Winnings cards
O 10 coins for risk perception questions

ASSIGN NUMBER TO FARMER → WRITE ON WINNINGSCARD, SMALL CARD AND QUESTIONNAIRE

ASSIGN ENUMERATOR TO FARMER

EXPLAIN WE PLAY TWO GAMES IN WHICH THEY CAN EARN MONEY. PAYOFF AT THE END WILL BE A DRAW FROM THESE TWO GAMES. MINIMUM 50 BIRR; MAXIMUM 160 BIRR

ENUMERATOR HAS BROWN ENVELOPE WITH 200 BIRR

EDUCATION: EXPERIENCING WEATHER AND CROP LOSS

TWO PRACTICE ROUNDS. FARMERS WHO GET BLUE GET 100 BIRR; YELLOW 28 BIRR.

EDUCATION RISK-SHARING

PLAY RISK-SHARING WITH ENUMERATORS AS EXAMPLE.

FARMERS WILL BE TEAMED UP (LIKE THE ENUMERATORS) WITH SOMEONE IN THE ROOM.

AT END RECOLLECT 200 BIRR FROM FARMERS

T3; rainfall insurance; farmer i and j

FARMER HAS EMPTY ENVELOPPE. TELL FARMER WINNINGS WILL BE COLLECTED IN ENVELOPPE.
MAKE SURE YOU HAVE ENVELOPE WITH 200 BIRR
PLAY RISK-SHARING WITHOUT RAINFALL INSURANCE WITH ENUMERATORS AS EXAMPLE
EDUCATION: EXPERIENCING RAINFALL INSURANCE
PLAY RISK-SHARING WITH RAINFALL INSURANCE WITH ENUMERATORS AS EXAMPLE

ASK UNDERSTANDING QUESTIONS:

ET3.1 If farmer rolls a yellow dice (bad crop) how much is the value of his crops in birr? …………..BIRR (right answer = 28)
ET3.2 If crop is bad (yellow dice) but weather is good (blue token)→ How much does farmer think is …….with rainfall insurance:
- His/her insurance claim? O 0 birr O 72 birr (right answer=0)
- His/her eventual payoff? …………BIRR (right answer= 28 – 18 = 10)

ET3.3 If weather is bad (yellow token) and crop is good (blue dice) → How much does farmer think is …….with rainfall insurance:
- His/her insurance claim O 0 birr O 72 birr (right answer= 72)
- His/her eventual payoff ………..BIRR (right answer = 100 – 18 + 72 = 154)

ET3.4 If weather is good and crop is good → How much does farmer think is …….with rainfall insurance:
- His/her insurance claim O 0 birr O 72 birr (right answer= 0)

- His/her eventual payoff ………..BIRR (right answer = 100 – 18 = 82)
ET3.5 If weather is bad and crop is bad → How much does farmer think is …….with rainfall insurance:
- His/her insurance claim O 0 birr O 72 birr (right answer= 72)
- His/her eventual payoff ………..BIRR (right answer = 28-18 + 72 = 82)
ET3.6 How much is the insurance premium? ………Birr (right answer = 18 Birr)

WE ARE NOW GOING TO PLAY THE REAL GAME
TELL FARMER IF HE/SHE IS FARMER i or FARMER j O i O j

GIVE FARMER 18 BIRR TO PUT IN WHITE ENVELOPE
TELL FARMER IF HE/SHE IS TEAMED-UP WITH j FROM O same iddir O different iddir

T3; farmer j; TREATMENT RAINFALL INSURANCE AND RISK-SHARING

TELL FARMER THAT IF HE/SHE WANTS TO TAKE INSURANCE THE PREMIUM (18 BIRR) WILL BE DEDUCTED FROM THEIR PAYOFF AT THE END (100-18 OR 28-18 OR 100-18+72 OR 28-18+72 + TRANSFER).

SHOW ON RISK-SHARING SHEET

SESSION LEADER EXPLAINS

T3j.1Does farmer want to take insurance? O yes O no **[RECORD ON WINNINGSCARD]**

LET THE FARMER PAY THE PREMIUM OF 18 BIRR IN CASE YES.

T3j.2 Weather (token draw) O drought (yellow) O medium/good weather (blue) **[RECORD ON WINNINGSCARD]**

PAY FARMER i 100 BIRR

T3j.3 Farmer dice roll O 100 BIRR (good crop) O 28 BIRR (bad crop)

GIVE FARMER 100 or 28 BIRR

PAY INSURANCE PAYOUT (72 BIRR)TO FARMER IF TOKEN IS YELLOW.

[RECORD ON WINNINGSCARD]

[BRING WINNINGSCARD TO SESSION LEADER]

T3; farmer i; TREATMENT RAINFALL INSURANCE AND RISK-SHARING

(J WILL DECIDE IF HE/SHE WANTS TO TAKE INSURANCE FIRST)

- REMIND THE FARMER THAT THE RAINFALL INSURANCE DOES NOT PAY FARMER j IF WEATHER IS GOOD BUT FARMER HAS BAD CROP → IF FARMER HAS BAD CROP HE WILL HAVE 10 BIRR 3/5 and 82 BIRR 2/5 times.
- REMIND FARMER THAT j MAY ALSO GET 154 WITH GOOD CROP AND RAINFALL INSURANCE PAYOUT.

T3i.1 Weather (token draw) O drought (yellow) O medium/good weather (blue)

PAY FARMER i 100 BIRR FROM BROWN ENVELOPPE

REMIND FARMER THAT J MAY EXPERIENCE A LOSS.

T3i.2 **(READ NOTE TO ENUMERATOR FIRST)** How much does i want to share with j in case j has bad crop? ……………BIRR **[RECORD ON WINNINGSCARD]**

NOTE TO ENUMERATOR: SHOW ON RISK-SHARING SHEET (BASED ON TOKEN DRAWN) THE TWO POSSIBILITIES:

- IF TOKEN DRAWN BLUE → j HAS EITHER 10 BIRR IF CHOSEN INSURANCE OR 28 BIRR IF CHOSEN NO INSURANCE
- IF TOKEN DRAWN YELLOW → j HAS EITHER 82 BIRR IF CHOSEN INSURANCE OR 28 BIRR IF CHOSEN NO INSURANCE

REMIND FARMER i THAT j WOULD GET 154 BIRR IN CASE OF TOKEN YELLOW BUT NO CROP LOSS → SHOW ON RISK-SHARING SHEET

T3i.3 **(READ NOTE TO ENUMERATOR FIRST)** How much does i want to share with j in case j has bad crop and took insurance?...........BIRR

NOTE TO ENUMERATOR: SHOW ON RISK-SHARING SHEET (BASED ON TOKEN DRAWN) THE PAYOFF OF j (if j took insurance and has shock j will have 70 birr)

T3i.3A What best describes this choice? O compensate because bad crop O compensate less because insurance O compensate more because this insurance doesn't fully protect O punish because take insurance and this insurance is risky O reward because take insurance to protect

T3i.4 (READ **NOTE TO ENUMERATOR** BEFORE ASKING QUESTION) How much does i want to share with j in case j has bad crop and took no insurance?...........BIRR

NOTE TO ENUMERATOR: SHOW ON RISK-SHARING SHEET (BASED ON TOKEN DRAWN) THE PAYOFF OF j (if j did not take insurance and has shock j will have 28 birr)

T3i.4A What best describes this choice? O compensate because bad crop O compensate more because no insurance O compensate more because this insurance doesn't fully protect
O punish because did not take insurance

T3i.5 How likely, on a scale from 0-10 (10 likely, 0 unlikely) does farmer i think it is that farmer j will take insurance?

O 0 O1 O2 O3 O4 O5 O6 O7 O8 O9 O10

RISK PREFERENCES GAME 2:

*******BEFORE EVERY GAME: REMIND FARMER THAT PAYOFF AT END DEPENDS ON OUTCOMES IN THIS GAME!!!!!********

G1.1	#	L	S	Sure	G1.2	#	L	S	Sure	G1.4	#	L	S	Sure
Lottery	1	O	O	0	Lottery	1	O	O	0	Lottery	1	O	O	0
	2	O	O	10		2	O	O	10		2	O	O	10
Win 160, p=0.5	3	O	O	20	Win 160, p=0.25	3	O	O	20	Win 160, p=0.75	3	O	O	20
	4	O	O	30		4	O	O	30		4	O	O	30
	5	O	O	40		5	O	O	40		5	O	O	40
**********	6	O	O	50	*******	6	O	O	50	***	6	O	O	50
REMIND FARMER THAT PAYOFF AT END DEPENDS ON OUTCOMES IN THIS GAME!!!!!	7	O	O	60	**REMIND FARMER THAT PAYOFF AT END DEPENDS ON OUTCOMES IN THIS GAME!!!!!**	7	O	O	60	**REMIND FARMER THAT PAYOFF AT END DEPENDS ON OUTCOMES IN THIS GAME!!!!!**	7	O	O	60
	8	O	O	70		8	O	O	70		8	O	O	70
	9	O	O	80		9	O	O	80		9	O	O	80
	10	O	O	90		10	O	O	90		10	O	O	90
	11	O	O	100		11	O	O	100		11	O	O	100
	12	O	O	110		12	O	O	110		12	O	O	110
	13	O	O	120		13	O	O	120		13	O	O	120
	14	O	O	130		14	O	O	130		14	O	O	130
	15	O	O	140		15	O	O	140		15	O	O	140
	16	O	O	150		16	O	O	150		16	O	O	150
	17	O	O	160		17	O	O	160		17	O	O	160

*******BEFORE EVERY GAME: REMIND FARMER THAT PAYOFF AT END DEPENDS ON OUTCOMES IN THIS GAME!!!!!********

G1.6	#	L	S	Sure	G1.7	#	L	S	Sure	G1.9	#	L	S	Sure
Lottery	1	O	O	-160	Lottery	1	O	O	-160	Lottery	1	O	O	-160
	2	O	O	-150		2	O	O	-150		2	O	O	-150
Lose 160, P=0.5	3	O	O	-140	Lose 160, p=0.25	3	O	O	-140	Lose 160, P=0.75	3	O	O	-140
	4	O	O	-130		4	O	O	-130		4	O	O	-130
	5	O	O	-120		5	O	O	-120		5	O	O	-120
****	6	O	O	-110	********	6	O	O	-110	*******	6	O	O	-110
REMIND FARMER THAT PAYOFF AT END DEPENDS ON OUTCOMES IN THIS GAME!!!!!	7	O	O	-100	**REMIND FARMER THAT PAYOFF AT END DEPENDS ON OUTCOMES IN THIS GAME!!!!!**	7	O	O	-100	**REMIND FARMER THAT PAYOFF AT END DEPENDS ON OUTCOMES IN THIS GAME!!!!!**	7	O	O	-100
	8	O	O	-90		8	O	O	-90		8	O	O	-90
	9	O	O	-80		9	O	O	-80		9	O	O	-80
	10	O	O	-70		10	O	O	-70		10	O	O	-70
	11	O	O	-60		11	O	O	-60		11	O	O	-60
	12	O	O	-50		12	O	O	-50		12	O	O	-50
	13	O	O	-40		13	O	O	-40		13	O	O	-40
	14	O	O	-30		14	O	O	-30		14	O	O	-30
	15	O	O	-20		15	O	O	-20		15	O	O	-20
	16	O	O	-10		16	O	O	-10		16	O	O	-10
	17	O	O	0		17	O	O	0		17	O	O	0

G1.11	blue	Yellow	#	L	S	**Sure**
Mixed Lottery:	160	-160	1	O	O	**0**
Win 160 if blue,	160	-150	2	O	O	
or lose an increasing	160	-140	3	O	O	
amount if yellow	160	-130	4	O	O	
with p=0.5	160	-120	5	O	O	
	160	-110	6	O	O	
	160	-100	7	O	O	
	160	-90	8	O	O	
********REMIND FARMER**	160	-80	9	O	O	
THAT PAYOFF AT END	160	-70	10	O	O	
DEPENDS ON OUTCOMES	160	-60	11	O	O	
IN THIS GAME!!!!!	160	-50	12	O	O	
	160	-40	13	O	O	
	160	-30	14	O	O	
	160	-20	15	O	O	
	160	-10	16	O	O	
	160	0	17	O	O	

Binswanger O A O B O C O D O E O F

3. DETERMINING PAYOFF GAME 2:

3.1 Random draw of one of the 11 price lists (G1.1-G1.11)
O 1 O 2 O 3 O 4 O 5 O 6 O 7 O 8 O 9 O 10 O 11
3.2 Random draw of one of the 17 decision rows decision row: .....
3.3 Look up farmer's choice in drawn price list/row O lottery O sure
if subject chose lottery:
3.4 put tokens price list and draw token: O Blue O Yellow

FILL OUT WINNINGSCARD AND GIVE TO SESSION LEADER

NOW START WORKING ON PDA

Risk-sharing Questionnaire T4

Date:

Session:

Session leader

Name enumerator:

Farmer number:

Before starting check (tick box on all):
O 1 questionnaire
O 1 red dice; 1 white dice in cup
O 1 white envelope → empty
O 1 brown envelope → 200 birr
O PDA
O 'Risk-sharing sheets'
O 'Risk-preference sheets'
O Pencil
O Winnings cards
O 10 coins for risk perception questions

ASSIGN NUMBER TO FARMER → WRITE ON WINNINGSCARD, SMALL CARD AND QUESTIONNAIRE
ASSIGN ENUMERATOR TO FARMER
EXPLAIN WE PLAY TWO GAMES IN WHICH THEY CAN EARN MONEY. PAYOFF AT THE END WILL BE A DRAW FROM THESE TWO GAMES. MINIMUM 50 BIRR; MAXIMUM 160 BIRR
ENUMERATOR HAS BROWN ENVELOPE WITH 200 BIRR
EDUCATION: EXPERIENCING WEATHER AND CROP LOSS
TWO PRACTICE ROUNDS. FARMERS WHO GET BLUE GET 100 BIRR; YELLOW 28 BIRR.
EDUCATION RISK-SHARING
PLAY RISK-SHARING WITH ENUMERATORS AS EXAMPLE.
FARMERS WILL BE TEAMED UP (LIKE THE ENUMERATORS) WITH SOMEONE IN THE ROOM.
AT END RECOLLECT 200 BIRR FROM FARMERS

T4; crop and insurance; farmer i and j

FARMER HAS EMPTY ENVELOPPE. TELL FARMER WINNINGS WILL BE COLLECTED IN ENVELOPPE.

MAKE SURE YOU HAVE ENVELOPE WITH 200 BIRR
PLAY RISK-SHARING WITHOUT ANY INSURANCE WITH ENUMERATORS AS EXAMPLE
EDUCATION: EXPERIENCING CROP INSURANC AND RAINFALL INSURANCE
PLAY RISK-SHARING WITH CROP INSURANCE AND RAINFALL INSURANCE WITH ENUMERATORS AS EXAMPLE → explain either no insurance offer, or indemnity insurance offer, or index insurance offer.

ASK UNDERSTANDING QUESTIONS:
ET4.1 If farmer rolls a yellow dice (bad crop) how much is the value of his crops in birr? (right answer = 28)
ET4.2 If crop is bad (yellow dice) but weather is good (blue token)→ How much does farmer think is …….with rainfall insurance:
- His/her insurance claim? O 0 birr O 72 birr (right answer=0)
- His/her eventual payoff? …………BIRR (right answer= 28 – 18 = 10)

ET4.3 If weather is bad (yellow token) and crop is good (blue dice) → How much does farmer think is …….with crop insurance:
- His/her insurance claim O 0 birr O 72 birr (right answer=0)
- His/her eventual payoff ………..BIRR (right answer = 100 – 30 = 70)

ET4.4 If weather is bad (yellow token) and crop is good (blue dice) → How much does farmer think is …….with rainfall insurance:
- His/her insurance claim O 0 birr O 72 birr (right answer= 72)
- His/her eventual payoff ………..BIRR (right answer = 100 – 18 + 72 = 154)

ET4.5 How much is the rainfall insurance premium?........BIRR (righanswer = 18 Birr)
ET4.6 How much is the crop insurance premium?........BIRR (righanswer = 30 Birr)

WE ARE NOW GOING TO PLAY THE REAL GAME
TELL FARMER IF HE/SHE IS FARMER i or FARMER j O i O j
TELL FARMER IF HE/SHE IS TEAMED-UP WITH j FROM O same iddir O different iddir

T4; farmer j; TREATMENT CROP OR RAINFALL INSURANCE RISK-SHARING

[FARMERS GET NOTE EITHER: NO INSURANCE; CROP INSURANCE; RAINFALL INSURANCE]

GIVE FARMERS EITHER 0; 18 or 30 BIRR.

TELL FARMER THAT IF HE/SHE WANTS TO TAKE INSURANCE THE PREMIUM (18 or 30 BIRR) WILL BE DEDUCTED FROM THEIR PAYOFF AT THE END (100-30 OR 28-30+72 OR 100-18 OR 28-18 OR 100-18+72 OR 28-18+72 + TRANSFER) → SHOW ON RISK-SHARING SHEET

LET THE FARMER PAY THE PREMIUM OF 18 OR 30 BIRR IN CASE YES

ANSWER EITHER T4j.1A; T4j.1B or T4j.1C

T4j.1A Tick if farmer has 'NO INSURANCE' → O 'NO INSURANCE'

T4j.1b Does farmer want to take crop insurance? O yes O no

- OR ASK FARMER:

T4j.1B Does farmer want to take rainfall insurance? O yes O no

- REMIND THE FARMER THAT THE RAINFALL INSURANCE DOES NOT PAY FARMER j IF WEATHER IS GOOD BUT FARMER HAS BAD CROP → IF FARMER HAS BAD CROP HE WILL HAVE 10 BIRR 3/5 and 82 BIRR 2/5 times.
- REMIND FARMER THAT j MAY ALSO GET 154 WITH GOOD CROP AND RAINFALL INSURANCE PAYOUT.

T4j.2 Weather (token draw) O drought (yellow) O medium/good weather (blue) **[RECORD ON WINNINGSCARD]**

PAY FARMER i 100 BIRR

T4j.3 Dice roll O drought (yellow) O medium/good weather (blue) **[RECORD ON WINNINGSCARD]**

T4j.4 Farmer pay-out: O yes O no

T4; farmer i; TREATMENT RAINFALL INSURANCE AND RISK-SHARING

J IS OFFERED EITHER CROP OR RAINFALL INSURANCE.

J WILL DECIDE IF HE/SHE WANTS TO TAKE INSURANCE FIRST.

- REMIND THE FARMER THAT THE RAINFALL INSURANCE DOES NOT PAY FARMER j IF WEATHER IS GOOD BUT FARMER HAS BAD CROP (10 BIRR) BUT j MAY ALSO GAMBLE AND

HOPE HE GETS GOOD CROP EVEN IF WEATHER IS BAD AND GET PAYOUT OF 154 BIRR. SO ACTUALLY THE FARMER IS TAKING RISK WITH RAINFALL INSURANCE.

T4i.1 Weather (token draw) O drought (yellow) O medium/good weather (blue)

- REMIND FARMER THAT HE WILL GET 100 BIRR FOR SURE THIS SEASON BUT FARMER j MAY EXPERIENCE A LOSS.

T4i.2 **(READ NOTE TO ENUMERATOR BEFORE ASKING QUESTION)** How much does i want to share with j in case j has bad crop? ……………BIRR **[RECORD ON WINNINGSCARD]**

NOTE TO ENUMERATOR: SHOW ON RISK-SHARING SHEETs (BASED ON TOKEN DRAWN) THE 6 POSSIBILITIES:

IF TOKEN DRAWN BLUE →

- 'NO INSURANCE' → IN BOTH CASES 28 BIRR
- WITH RAINFALL → j HAS EITHER 10 BIRR IF RAINFALL INSURANCE OR 28 BIRR IF CHOSEN NO RAINFALL INSURANCE
- WITH CROP → j HAS EITHER 70 BIRR IF CHOSEN CROP INSURANCE OR 28 BIRR IF CHOSEN NO CROP INSURANCE

IF TOKEN DRAWN YELLOW →

- 'NO INSURANCE' → IN BOTH CASES 28 BIRR
- WITH RAINFALL → j HAS EITHER 82 BIRR IF CHOSEN RAINFALL INSURANCE OR 28 BIRR IF CHOSEN NO RAINFALL INSURANCE
- WITH CROP → j HAS EITHER 70 BIRR IF CHOSEN CROP INSURANCE OR 28 BIRR IF CHOSEN NO CROP INSURANCE

REMIND FARMER i THAT j WOULD GET 154 BIRR IN CASE OF TOKEN YELLOW, RAINFALL INSURANCE BUT NO CROP LOSS → SHOW ON RISK-SHARING SHEET

T4I.3 **(FARMER OFFERED NO INSURANCE):** How much does i want to share with j in case j has bad crop(i has 100 BIRR, j has 28 BIRR)?........BIRR

T4i.4 **READ NOTE TO ENUMERATOR BEFORE ASKING QUESTION: (FARMER OFFERED CROP INSURANCE AND TOOK CROP INSURANCE)** How much does i want to share with j in case j has bad crop and took crop insurance?...........BIRR

NOTE TO ENUMERATOR: SHOW ON RISK-SHARING SHEET (BASED ON TOKEN DRAWN) THE PAYOFF OF j (if j took crop insurance and has shock j will have 70 birr)

T4i.4A What best describes this choice? O compensate because bad crop O compensate less because insurance O punish because take insurance O reward because take insurance

T4i.5 **READ NOTE TO ENUMERATOR BEFORE ASKING QUESTION: (FARMER OFFERED CROP INSURANCE AND DID NOT TAKE CROP INSURANCE)** How much does i want to share with j in case j has bad crop and took no crop insurance?...........BIRR

NOTE TO ENUMERATOR: SHOW ON RISK-SHARING SHEET (BASED ON TOKEN DRAWN) THE PAYOFF OF j (if j took no crop insurance and has shock j will have 28 birr)

T2i.4A What best describes this choice? O compensate because bad crop O punish because not take insurance

T4i.6 **READ NOTE TO ENUMERATOR BEFORE ASKING QUESTION: (FARMER OFFERED RAINFALL INSURANCE AND TOOK RAINFALL INSURANCE)** How much does i want to share with j in case j has bad crop and took rainfall insurance?...........BIRR

NOTE TO ENUMERATOR: SHOW ON RISK-SHARING SHEET (BASED ON TOKEN DRAWN) THE PAYOFF OF j (if j took insurance and has shock j will have 10 birr)

T3i.3A What best describes this choice? O compensate because bad crop O compensate less because insurance O compensate more because this insurance doesn't fully protect O punish because take insurance and this insurance is risky O reward because take insurance to protect

T4i.7 **READ NOTE TO ENUMERATOR BEFORE ASKING QUESTION: (FARMER OFFERED RAINFALL INSURANCE AND DIDN'T TAKE RAINFALL INSURANCE)** How much does i want to share with j in case j has bad crop and took no rainfall insurance?...........BIRR

NOTE TO ENUMERATOR: SHOW ON RISK-SHARING SHEET (BASED ON TOKEN DRAWN) THE PAYOFF OF j (if j took no rainfall insurance and has shock j will have 28 birr)

T3i.4A What best describes this choice? O compensate because bad crop O compensate more because no insurance O compensate more because this insurance doesn't fully protect
O punish because did not take insurance

T4i.8 If farmer j was offered rainfall insurance, how likely, on a scale from 0-10 (10 likely, 0 unlikely) does farmer i think it is that farmer j will have taken rainfall insurance?

O 0 O1 O2 O3 O4 O5 O6 O7 O8 O9 O10

T4i.9 If farmer j was offered crop insurance, how likely, on a scale from 0-10 (10 likely, 0 unlikely) does farmer i think it is that farmer j will have taken crop insurance?

O 0 O1 O2 O3 O4 O5 O6 O7 O8 O9 O10

RISK PREFERENCES GAME 2:

*******BEFORE EVERY GAME: REMIND FARMER THAT PAYOFF AT END DEPENDS ON OUTCOMES IN THIS GAME!!!!!********

G1.1	#	L	S	Sure	G1.2	#	L	S	Sure	G1.4	#	L	S	Sure
Lottery Win 160, p=0.5 ********** **REMIND FARMER THAT PAYOFF AT END DEPENDS ON OUTCOMES IN THIS GAME!!!!!**	1	O	O	0	Lottery Win 160, p=0.25 ******* **REMIND FARMER THAT PAYOFF AT END DEPENDS ON OUTCOMES IN THIS GAME!!!!!**	1	O	O	0	Lottery Win 160, p=0.75 *** **REMIND FARMER THAT PAYOFF AT END DEPENDS ON OUTCOMES IN THIS GAME!!!!!**	1	O	O	0
	2	O	O	10		2	O	O	10		2	O	O	10
	3	O	O	20		3	O	O	20		3	O	O	20
	4	O	O	30		4	O	O	30		4	O	O	30
	5	O	O	40		5	O	O	40		5	O	O	40
	6	O	O	50		6	O	O	50		6	O	O	50
	7	O	O	60		7	O	O	60		7	O	O	60
	8	O	O	70		8	O	O	70		8	O	O	70
	9	O	O	80		9	O	O	80		9	O	O	80
	10	O	O	90		10	O	O	90		10	O	O	90
	11	O	O	100		11	O	O	100		11	O	O	100
	12	O	O	110		12	O	O	110		12	O	O	110
	13	O	O	120		13	O	O	120		13	O	O	120
	14	O	O	130		14	O	O	130		14	O	O	130
	15	O	O	140		15	O	O	140		15	O	O	140
	16	O	O	150		16	O	O	150		16	O	O	150
	17	O	O	160		17	O	O	160		17	O	O	160

*******BEFORE EVERY GAME: REMIND FARMER THAT PAYOFF AT END DEPENDS ON OUTCOMES IN THIS GAME!!!!!********

G1.6	#	L	S	Sure	G1.7	#	L	S	Sure	G1.9	#	L	S	Sure
Lottery Lose 160, P=0.5 **** **REMIND FARMER THAT PAYOFF AT END DEPENDS ON OUTCOMES IN THIS GAME!!!!!**	1	O	O	-160	Lottery Lose 160, p=0.25 ******** **REMIND FARMER THAT PAYOFF AT END DEPENDS ON OUTCOMES IN THIS GAME!!!!!**	1	O	O	-160	Lottery Lose 160, P=0.75 ******* **REMIND FARMER THAT PAYOFF AT END DEPENDS ON OUTCOMES IN THIS GAME!!!!!**	1	O	O	-160
	2	O	O	-150		2	O	O	-150		2	O	O	-150
	3	O	O	-140		3	O	O	-140		3	O	O	-140
	4	O	O	-130		4	O	O	-130		4	O	O	-130
	5	O	O	-120		5	O	O	-120		5	O	O	-120
	6	O	O	-110		6	O	O	-110		6	O	O	-110
	7	O	O	-100		7	O	O	-100		7	O	O	-100
	8	O	O	-90		8	O	O	-90		8	O	O	-90
	9	O	O	-80		9	O	O	-80		9	O	O	-80
	10	O	O	-70		10	O	O	-70		10	O	O	-70
	11	O	O	-60		11	O	O	-60		11	O	O	-60
	12	O	O	-50		12	O	O	-50		12	O	O	-50
	13	O	O	-40		13	O	O	-40		13	O	O	-40
	14	O	O	-30		14	O	O	-30		14	O	O	-30
	15	O	O	-20		15	O	O	-20		15	O	O	-20
	16	O	O	-10		16	O	O	-10		16	O	O	-10
	17	O	O	0		17	O	O	0		17	O	O	0

G1.11	blue	Yellow	#	L	S	**Sure**
Mixed Lottery:	160	-160	1	O	O	**0**
Win 160 if blue,	160	-150	2	O	O	
or lose an increasing	160	-140	3	O	O	
amount if yellow	160	-130	4	O	O	
with p=0.5	160	-120	5	O	O	
	160	-110	6	O	O	
	160	-100	7	O	O	
********REMIND FARMER**	160	-90	8	O	O	
THAT PAYOFF AT END	160	-80	9	O	O	
DEPENDS ON OUTCOMES	160	-70	10	O	O	
IN THIS GAME!!!!!	160	-60	11	O	O	
	160	-50	12	O	O	
	160	-40	13	O	O	
	160	-30	14	O	O	
	160	-20	15	O	O	
	160	-10	16	O	O	
	160	0	17	O	O	

Binswanger O A O B O C O D O E O F

3. DETERMINING PAYOFF GAME 2:

3.1 Random draw of one of the 11 price lists (G1.1-G1.11)
O 1 O 2 O 3 O 4 O 5 O 6 O 7 O 8 O 9 O 10 O 11
3.2 Random draw of one of the 17 decision rows decision row: .....
3.3 Look up farmer's choice in drawn price list/row O lottery O sure
if subject chose lottery:
3.4 put tokens price list and draw token: O Blue O Yellow

FILL OUT WINNINGSCARD AND GIVE TO SESSION LEADER

NOW START WORKING ON PDA