The London School of Economics

and Political Science

Four Essays in Agricultural and Development Economics

Xavier Vollenweider

A thesis submitted to the Department of Geography and Environment of the London School of Economics for the degree of Doctor of Philosophy, London, February 2016

Declaration

I certify that the thesis I have presented for examination for the MPhil/PhD degree of the London School of Economics and Political Science is solely my own work other than where I have clearly indicated that it is the work of others (in which case the extent of any work carried out jointly by me and any other person is clearly identified in it).

I warrant that this authorisation does not, to the best of my belief, infringe the rights of any third party.

I declare that my thesis consists of 60’000 words.

Chapter 3 is co-authored with Erwin Bulte, Salvatore Di Falco and Menale Kassie. Erwin Bulte and Salvatore Di Falco provided the experimental design, Menale Kassie selected the seeds varieties and the sample of participants in Tanzania, I provided the theoretical model to interpret the results, the empirical strategy, did the estimation, and supervised the data collection. The introduction of the chapter was co-written with Salvatore Di Falco, Menale Kassie contributed also to the data section.

Chapter 4, based on the experiment cited above, has been co-authored with Salvatore Di Falco. We contributed equally to the empirical section and to the introduction. I provided the theoretical model and did the estimation.

A draft of chapters numbered 2 and 4 of this thesis were copy edited for conventions of language, spelling and grammar by Grove Proofreading & Editorial Services and chapter numbered 1 to 5 by Barry Monaghan.

To Laura

Acknowledgments

I would like to first extend my thanks to my supervisors Salvatore Di Falco, Steve Gibbons and Cathal O'Donogue for their guidance over the course of the PhD. I would like also to extend my thanks to Ben Groom and Kelvin Balcombe for their thorough and constructive review of this thesis. I also received very helpful comments from Elvezio Ronchetti. My thoughts go also to Ruth Madulu and Theresia Gregory for their professionalism and expertise in managing the randomized controlled trials in Tanzania and to the farmers who took part to it and provided hospitality during the field visits. Barry Monaghan provided an invaluable help by correcting grammatical errors and other linguistic mischief. I would like to also thank my parents for their constant encouragement. I feel privileged to have formed friendships with Misato Sato, Alessandro Tavoni, Antoine Dechezleprêtre, Raphael Calel, Chris Gaskell, Charlene Watson and Murray Collins at LSE. I will never forget our time together in London. My thanks go also to my brother Louis and all my friends in Geneva. My deepest gratitude and my love go to Laura Platchkova, my source of joy and motivation.

This work was supported by the Walsh fellowship and the Rural Economy and Development Programme of Teagasc, the Agriculture and Food Development Authority of the Republic of Ireland. I thank also the Grantham research institute and the Teagasc Rural Economy Research centre for hosting as well as the Selian Agricultural Research Institute in Arusha for providing help in conducting the experiment in Tanzania.

Abstract

In the first paper, I introduce a new framework to estimate household climate risk exposure based on a combination of climate and microeconomic data. I apply it to the Ethiopian Rural Household Survey (1994-2009) and find that households living at low altitudes are the most vulnerable to weather shocks. The second paper is based on a combination of open and double-blind randomized controlled trials (RCT) conducted in Tanzania in 2013 with 560 farmers. By comparing the results between the participants in the open and double-blind groups, we find that more than 50% of the total effect of improved seeds estimated in traditional open RCTs depends on farmers' behaviour. The third paper, based on the RCT mentioned above (only the open one is used), tests the hypothesis that farmers try to escape forced solidarity when facing favourable conditions. We find that farmers having received the improved seeds decrease their number of social interactions. We interpret this as a sign that farmers seek to hide from the pressure to redistribute. In the fourth paper, I leave Africa for the Republic of Ireland and show that a large Irish agri-environmental scheme does not increase farmers' risk exposure.

Table of Contents

List of figures

List of tables

Summary

According to the median estimates of the United Nations Population Division (United Nations 2015), the world population will reach 9.7 billion in 2050. In order to feed this larger, richer and more urban population, food production will have to increase by 70% above 2009 levels (net of bio-fuels use) while livestock demand could already be up by 68% by 2030 (FAO 2009). Assuming that the goal is to meet this challenge with no agricultural land expansion in order to limit deforestation, currently responsible for 17% of greenhouse gas emissions (IPCC 2007), a yield growth of 1.07% per annum (p.a.) is required globally, while this figure goes up to 4.6% p.a. for Africa. However, yield growth has declined over the last decades and is projected to be around 0.7% p.a. (half their historical trend) for the decades to come (FAO 2009). Furthermore, it is very likely that climate change will cause, by 2050 in Sub-Saharan Africa, a 7% yield loss for the main food crops while its impact on rain-fed agriculture could be even more severe (up to 50% in some areas, IPCC 2007). There is hence a need to increase the yield of food production, to decrease its vulnerability to climate risk while limiting its impact on the environment. This tryptic has been labelled as the need a for a 'doubly-green revolution' (Conway 1998).

The first chapter provides an overview of the literature on the link between risk, poverty, and agricultural technology adoption with an emphasis on empirical studies conducted in Sub-Saharan Africa. I start with an introduction to the expected utility theory and its limitations. The classic risk estimation framework in agricultural economics and recent developments are then discussed. The rest of the review surveys empirical studies on the link between risk, poverty and technology adoption. Social networks in developing countries play an important role in technology diffusion and household risk-management strategies. They are hence discussed in corresponding sections. We conclude by highlighting the contribution of the thesis to some of the topics introduced in this literature review.

In the second chapter, we introduce a new framework to estimate climate risk exposure with the standardized precipitation evapotranspiration index (SPEI) as its building block. The approach is simple enough to accommodate quantile regressions and hence offer the opportunity to broaden the scope of the analysis to different categories of the population. The main contribution of the chapter is to provide an estimation framework where the various measure of risk, such as variance and skewness, can be directly derived from the regression parameters estimates of the SPEI. The methodology is illustrated with a case study on Ethiopia. In accordance with previous studies, the results show that households located at low altitudes are the most exposed to climate risk.

In the third chapter, we assess the role of farmers' behaviour in driving the yield increase of improved maize seeds. The study is based on the combination of open and double-blind randomized controlled

trials (RCTs) conducted in 2013 in two regions of Tanzania with 560 farmers. The advantage of combining open and double-blind RCTs is to allow the distinction between the effects of the improved seeds *per se* from the effects resulting from a change in the management of the farm. The empirical contribution of this study is to show that this behavioural response plays a central role in driving the increase in yields brought about improved seeds. In our experiment, more than 50% of the increase in yield estimated in the traditional open RCT would not have materialized without the behavioural response.

Social networks play an important role in the livelihood of rural communities in Sub-Saharan Africa. The more successful members of the network must help the least successful or unlucky members of the social network. Recently, some observational and experimental evidence has indicated that these obligations may trigger an evasive response. In the fourth chapter, we investigate if participants to the RCT conducted in Tanzania try to escape forced solidarity when facing favourable conditions by hiding from their network. We find that farmers who were allocated improved seeds decreased the number of their social interactions, particularly if they have numerous relatives in the village. We interpret this as a sign that farmers attempt to escape forced solidarity and that the pressure to share increases as the size of the social network increases. To our knowledge, this constitutes the first set of evidence of evasive behaviour based on data from a RCT involving real interactions (i.e. not in a choice experiment or with observational data).

In the fifth chapter, we investigate the impact on Irish farmers' risk exposure of the Rural Environment Protection Scheme (REPS), an agri-environmental scheme. It has been shown in the literature that organic and, more generally, low input agriculture tends to increase risk exposure while risk aversion plays a role in the low adoption of sustainable production techniques. We show that REPS does not increase risk exposure, and adequately compensates farmers for foregone returns, which might be one of the reasons of its large success among farmers. Addressing risk considerations in policies aimed at making farmers eco-friendlier is an important dimension to the challenge of preserving both farmers' quality of life and the environment.

References

Conway, Gordon (1998), The doubly green revolution: food for all in the twenty-first century (Cornell University Press).

FAO, Food and Agriculture Organization of the United Nations (2009), 'How to Feed the World in 2050', (Rome).

Intergovernmental Panel on Climate Change (2007), IPCC Fourth Assessment Report (AR4)

United Nations, Department of Economic and Social Affairs, Population Division (2015), 'World Population Prospects: The 2015 Revision, Key Findings and Advance Tables.', Working Paper, ESA/P/WP.241.

1 Literature Review

Xavier Vollenweider

Abstract

This chapter provides an overview of the literature on the link between poverty, risk, and technology adoption. We start by an introduction to the expected utility theory and discuss briefly the violation of the independence axiom and ambiguity aversion. The classic risk estimation framework in agricultural economics and its recent developments are then discussed in the following section. The rest of the review surveys empirical studies on the link between risk, poverty, and technology adoption. Social networks in developing countries play a role both in risk-coping and technology diffusion. They are hence discussed in corresponding sections. We conclude by highlighting the contribution of the thesis to some of the topics introduced in this literature review.

1.1 Introduction

According to the median estimates of the United Nations Population Division (United Nations 2015), the world population will reach 9.7 billion in 2050. In order to feed this larger, richer and more urban population, food production will have to increase by 70% (net of bio-fuels use) above 2009 levels while livestock demand could be up by 68% already by 2030 (FAO 2009). Assuming that the goal is to meet this challenge with no agricultural land expansion in order to limit deforestation, currently responsible for 17% of greenhouse gas emissions (IPCC 2007), a yield growth of 1.07% per annum (p.a.) is required globally, while this figure goes up to 4.6% p.a. for Africa. However, yield growth has declined over the last decades and is projected to be around 0.7% p.a. (half their historical trend) for the decades to come (FAO 2009). Furthermore, it is very likely that climate change will cause, by 2050 in Sub-Saharan Africa, a 7% yield loss for the main food crops while its impact on rain-fed agriculture could be even more severe (up to 50% in some areas, IPCC 2007). There is hence a need to increase the yield of food production, to decrease its vulnerability to climate risk while limiting its impact on the environment. This tryptic has been labelled the need for a 'doubly-green revolution' (Conway 1998).

The present thesis addresses some of the methodological and empirical aspects of this challenge. Chapter 2 provides a simple framework for estimating climate risk and climate vulnerability with a case study on Ethiopia. Chapter 3 analyses the role of effort allocation in driving the productivity increase of improved seeds thanks to the combination of open and double-blind randomized controlled trials (RCT) conducted in Tanzania. The open RCT data are also used in chapter 4 in an analysis of evasive response to social pressure to share risk and income among kin in village economy. Lastly, chapter 5 studies the effect of an agri-environmental scheme on risk exposure of Irish farmers. In order to set the scene, we provide below a survey of the literature on risk and technology adoption in developing countries.

For the sake of concision, some important streams of research are left out of this review notwithstanding their contribution to the general debate, notably the Ricardian approach (Mendelsohn et al. 1994), because its assumption of a well-functioning property market is rarely met in developing countries (Di Falco et al. 2011); the stochastic budgeting approach, because it relies on scenario analysis and simulation, rather than on empirical evidence; and the state-contingent approach (e.g. Chambers and Quiggin 1998), because there are only a few empirical studies based on it and it is not used in the present thesis.

The seminal paper by Sandmo (1971) showing that risk leads to underinvestment and underproduction contributed to establishing the economics of production under uncertainty as an important research stream in economics, with agriculture as one of its favourite case studies. If production risk is a major topic in the agricultural economics literature, it is probably because ’the most singular aspect of agricultural production is its randomness’ (Chambers and Quiggin 1998). Disparities are large in terms of farms’ risk exposure between developing countries (wherein traditional farming practices based on rain-fed agriculture dominate), and developed countries (wherein high-tech farms supply a large part of the alimentary needs and agricultural subsidies seek to shelter producers from market instabilities). Despite this heterogeneity, farming systems share the common attribute of being at the nexus of the markets and the environment.

Admittedly there are different types of risk faced by a farm (Hardaker et al. 2004; Hazell 1992): production risk (e.g. pest and animal diseases, droughts and floods), market risk (e.g. inputs and outputs price volatility); resource risk (e.g. fertilizers, seeds and labour supply shocks), institutional risk (e.g. changes in policy), financial risk (e.g. changes in the interest charged on farm debt), personal risk (e.g. health issues, accidents), asset risk (theft or fire damage to buildings, machinery and livestock). While price risks are the dominant source of income shocks in OECD countries, production risk is of major concern among small-scale subsistence farmers in developing countries.

The common understanding of risk is often associated with negative impacts but rarely articulated in terms of probabilities (Fafchamps 2009). Knight (1921) distinguished *risk*, wherein a probability set can be assigned to the set of possible outcomes, from *uncertainty*, wherein no probabilities can be assigned. Savage (1954) largely discarded this distinction by grounding *subjective probability* as the relevant concept to model agent perception of risk. The agricultural economics literature has adopted the expected utility theory first formalised by Bernoulli in 1738, and applied to modern economics by von Neumann and Morgenstern (1944), as the backbone of its decision-theoretic framework. Although this approach was challenged as early as 1953 by Allais and more generally by the field of behavioural economics (e.g. 1979; Machina 1987; Tversky and Kahneman 1974), no other theoretical decision framework has really succeeded in superseding it in the agricultural and development economics literature. We will nevertheless review briefly some of the main departures from expected utility theory reported in choice experiments.

In section 1.2, we provide an introduction to the main decision theoretical framework of agricultural and development economics, the expected utility theory, and flesh out some of its limitations. Section 1.3 presents the main risk estimation framework and recent developments. Section 1.4 provides an overview of the empirical findings on the link between risk and poverty. Section 1.5 looks at the determinants of new technology adoption. Social networks in developing countries play a role both in risk-coping and in technology diffusion. They are hence discussed in sections 1.4.3 and 1.5. We comment briefly on the role of risk in agri-environmental policies of developed countries in section

1.4.2. Section 1.6 concludes and discusses the contribution of the present thesis to the various research streams presented in this literature review.

1.2 Expected utility theory and risk aversion

We present the expected utility framework by relying on a series of examples and keeping the mathematical presentation as limited as possible. The goal is to introduce the decision theoretical framework in which most agricultural and development studies related to risk are grounded. We also review at the end of the section some of its limitation identified in the field of non-expected utility theory.

Let us assume that an agent can choose *either* to toss a coin giving him a one in two chance of winning £10 *or* to be directly paid a sure amount of money. The amount of sure money he asks for can inform us on his degree of risk aversion. A risk-neutral agent will be indifferent between betting and receiving straight away the exact average of the gamble (£5). By contrast, a risk-averse agent might prefer £3 to the risk of having nothing, because, as Rothschild and Stiglitz (1971) put it, 'a bird in the hand is worth two in the bush'. The risk-seeking agent, by contrast, will require a higher sure amount of money (let us say £7), because he prefers to play for the chance of earning £10.

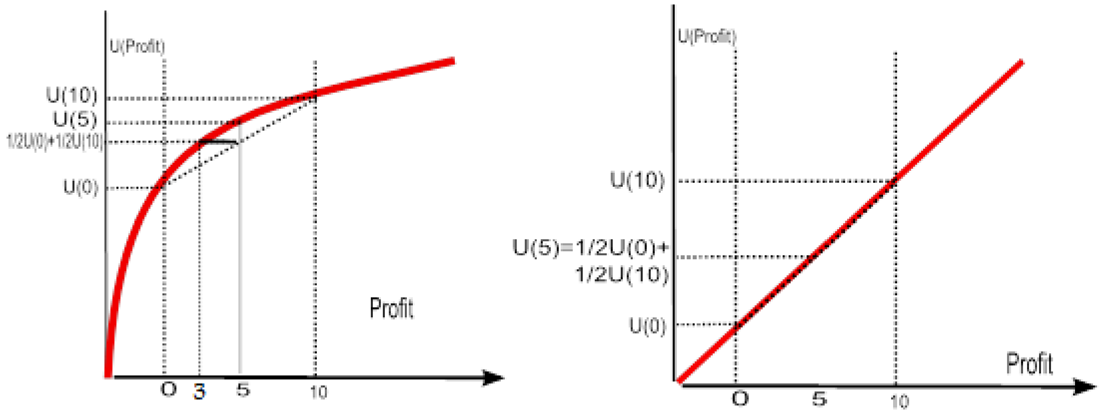
This sure amount of money is called the *certainty equivalent* (CE). We see that the smaller the CE is, the bigger the risk aversion. The difference between the expected gain (i.e. the average of the gamble) and the certainty equivalent is the *risk premium*: the risk-averse player has a risk premium of £2 (£5-£3), which is bigger than the risk neutral (£5-£5=£0) and risk-seeking one (£5-£7=-£2).

One can directly apply these insights to the interaction between risk aversion and agricultural production choices. Replacing the choice between coin toss and the sure sum of money by, respectively, the choice between planting a risky crop (e.g. coffee) or a riskless but low return crop (e.g. sweet potato), we can see that a risk-averse farmer will choose to plant sweet potatoes while a risk-seeking one will prefer planting coffee. The difference between the profit of the sweet potato and the coffee is the risk premium. Risk premium is also called the *private cost of risk bearing*, as it is the foregone returns paid for lower risk. This example suggests that risk can have an important cost for risk-averse farmers: to renounce more lucrative activities and, potentially, to be trapped in poverty.

From this simple example, we see that it is important to stress the difference between risk as a probability measure and the event to which these probabilities are attached, e.g. droughts or price collapse (Fafchamps 2009). Although the latter is directly tangible, the example illustrates that the risk in itself can constitute a constant draw on farmers' livelihoods, as they might adopt low-risk strategies at the cost of lower expected revenues.

Let us define these concepts a bit more formally. Figure 1.1 plots the utility function of a risk-averse (left) and a risk-neutral (right) agent defined in the domain of the coin toss example. For the risk-averse player, the sum of the utilities of the payoffs of the game, $1/2\,U(0) + 1/2U(10)$, is less than the utility of the average payoff of the game, $U(5)$. By contrast, both quantities are equal for the risk neutral agent. Hence his indifference towards playing, or directly receiving £5.

Figure 1.1: Utility function of a risk averse and risk neutral agent



The risk premium is represented on the left panel by the thick straight line under the utility curve between 3 and 5. It equals the average payoff of the game, £5, minus the CE payoff of the game, £3, which is £2. By contrast, we see that under risk neutrality, the utility function is linear and the CE payoff equals the average of the game, £5. The difference between these utility functions is their curvature which is captured by their second derivatives.

The Arrow-Pratt coefficient of absolute risk aversion, AP, (Pratt 1964) is based on the curvature of the utility function:

$$AP = -\frac{U''(\pi)}{U'(\pi)} \tag{1}$$

where π represents the payoffs of the game in our example. The second derivative of the utility function is normalized by the first derivative in order to offer comparable metrics at different level of wealth. Lastly, as U'' is negative, one adds a negative sign in order to end up with a more intuitive metric: a higher AP value means higher risk aversion.

To better understand the implication of choosing a particular functional form for the utility, let's specify it as an exponential function:

$$U(\pi) = 1 - \exp(-\alpha\pi) \tag{2}$$

Then the AP is:

$$AP = -\frac{U''(\pi)}{U'(\pi)} = -\left(\frac{-\alpha^2 e^{-\alpha\pi}}{\alpha e^{-\alpha\pi}}\right) = \alpha \tag{3}$$

We see that the level of absolute risk aversion, α, is independent of wealth and constant for the whole domain of the utility function. This means the agent exhibits constant absolute risk aversion (CARA). This implies that, under an exponential utility function specification, the multi-millionaire owner of a hacienda is as afraid of losing 1 dollar as landless farmers living on one dollar per day. Other utility functions allow more flexibility in terms of the structure of risk aversion.

Let us examine risk aversion in more detail by modifying the choice offered to our gamblers. Now, we ask them to choose between two different gambles: (1) if they score Heads, they receive £10, Tails £0, (2) if they score Heads, they receive £9, Tails £1. Although the expected gains are the same, the risk-averse player prefers the second gamble, because the spread between both payoffs is smaller, while the risk neutral player is indifferent. This aversion to the spread of payoffs is precisely what is captured by the Arrow-Pratt coefficient of absolute risk aversion, hence it is also called the coefficient of aversion to mean-preserving spread (Rothschild and Stiglitz 1970). In the case of agriculture, a farmer expecting a yield between 200 kg/ha and 4,000 kg/ha would have a higher spread than a farmer expecting a yield between 1,800 kg/ha and 2,200 kg/ha for instance. The spread is generally computed as the variance of the distribution of yield or output.

One peculiarity of betting with fair coins is that they are symmetrically distributed. However, payoffs linked to economic activities might not be so. The distribution of probabilities between good and bad events is also an important characteristic of risk. Let's take a simple example from the paper of Menezes et al. (1980) to illustrate this concept. Mao (1970) questioned executives in several industries about their preferences concerning two lotteries with payoffs x and corresponding probabilities $pr(x)$ described in (Table 1.1):

Table 1.1: Choice experiment on downside risk aversion

$f(x)$	$g(x)$
$Pr(x = 1) = 75\%$	$Pr(x = 0) = 25\%$
$Pr(x = 3) = 25\%$	$Pr(x = 2) = 75\%$

Mao (1970) found unambiguous preferences for $f(x)$ over $g(x)$ although these distributions have equal mean and variance, and despite the fact that the most likely outcome of $g(x)$, its mode, is higher. He interprets this preference as evidence of the importance of downside risk (here the risk of having 0) captured by the coefficient of asymmetry of payoffs, the skewness. In other terms, agents prefer positively skewed distribution to negatively skewed one because they dislike downside risk. Menezes (1980) and Eeckhoudt and Schlesinger (2005) show that the coefficient of downside risk aversion is given by:

$$DS = \frac{U'''(\pi)}{U'(\pi)} \tag{4}$$

In summary, we have stressed the differences between risk as a probability concept and the consequences that are attached to it. We then detailed two aspects of risk: the spread (variance) of the possible outcomes and their asymmetry (skewness, i.e. downside risk). The expected utility theory provides an elegant framework to represent the agents' preferences with respect to these various facets of risk.

Nevertheless, the expected utility rests on three assumptions: the ordering, continuity and independence of preferences with respect to the set of pairs of probability and payoffs (hereafter, prospects). The ordering axiom and continuity axiom imply that choice must be consistent, transitive. The evaluation of each prospect via the utility function can therefore be carried to any degree of precision, so that a certainty equivalent exists for each prospect (e.g Buschena 2003; Wakker 2010). The independence axiom implies that the choice between two pairs of prospects should not be affected if a third pair of prospects is mixed in with the first two pairs. The independence axiom was questioned as early as 1953 by Allais in choice-experiments. We follow below the presentation of Starmer (2000). Participants were offered the choice between the two lotteries presented in Table 1.2:

Table 1.2: Lottery 1, choice-experiment on the independence axiom

$f_1(x)$	$g_1(x)$
$P\,r(x = \text{£}1\text{m}) = 100\%$	$P\,r(x = \text{£}5\text{m}) = 10\%$
	$P\,r(x = \text{£}1\text{m}) = 89\%$
	$P\,r(x = \text{£}0) = 1\%$

Most of participants chose the lottery $f_1(x)$ as it provides £1 million with certainty. Participants were then asked to choose between the two lotteries presented in Table 1.3:

Table 1.3: Lottery 2, choice-experiment on the independence axiom

$f_2(x)$	$g_2(x)$
$Pr(x = £1\text{m}) = 11\%$	$Pr(x = £5\text{m}) = 10\%$
$Pr(x = £0\,) = 89\%$	$Pr(x = £0) = 90\%$

Most participants chose $g_2(x)$ because the chance of winning in $f_2(x)$ and $g_2(x)$ appear very similar while the prize of $g_2(x)$ is much higher. However, this constitutes a violation of the independence axiom.

Indeed, let us now rewrite the gamble $f_1(x) = \{£1m, 11\%; £1m, 89\%\}$ and $g_1 = \{£0, 1\%; £5, 10\%; £1m, 89\%\}$. Both $f_1(x)$ and $g_1(x)$ have ($£1m$, 89%) in common. Now, let rewrite $f_2(x) = \{£1m, 11\%; £0{,}89\%\}$ and $g_2(x) = \{£0{,}1\%; £5{,}10\%; £0{,}89\%\}$. Both $f_2(x)$ and $g_2(x)$ have as common consequence (£0,89%). The only change between the first and the second pairs of lotteries is the change in the common consequence from ($£1m$, 89%) to (£0,89%). Under the independence axiom, a change in the common consequence should not affect the choice between $\mathbf{f}$ and $\mathbf{g}$ lotteries. Many studies have confirmed that the choice between prospects was influenced by change in the common consequence (Starmer 2000)[1].

So far, the probabilities of each possibility where known. What happens when probabilistic judgments are hard to make? This is a situation which is much closer to decision-making in real conditions. For instance, a farmer will have only a vague idea of the weather conditions over the next three years when deciding to plant a coffee crop. As mentioned in the introduction, the distinction between risk, where probabilities can be attached to the events and uncertainty, where no probability are attached to the event, was largely discarded in favour of the notion of subjective probability proposed by Savage (Savage 1954). The decision under uncertainty can be reduced to decision under risk, with the assumption that agents hold subjective probabilistic beliefs which are used linearly.

However, Ellsberg (1961) showed that one could not discard the distinction between uncertainty and risk. The following though experiment will illustrate his point (see Etner et al. 2012). Say an individual faces two urns, with 100 balls in each urn: in urn I, he knows that there are 50 black and 50 red balls (he can open the urn and count them); in urn II, he does not know the proportion of black and red

[1] Another violation of the independence axiom is the common ratio effect (Starmer, 2000). Say you have to choose between the following lotteries: $f_1{}'(£3000{,}100\%; £0{,}1-100\%)$ and $g_1{}'(£4000{,}80\%; £0{,}1-80\%)$. Many people would go for $f_1{}'$ because of the certainty of getting £3000. Now say you have to choose between $f_2{}'(£3000{,}25\%; £0{,}1-25\%)$ and $g_2{}'(£4000{,}20\%; £0{,}1-20\%)$. Many people would go for $g_2{}'$. However, expected utility theory predicts that one would go either for $\boldsymbol{f}'$ or $\boldsymbol{g}'$ gambles as the chance of winning as only been divided by 4 in both lotteries.

balls. He can choose a colour and then pick at random one ball, but he has first to decide from which urn he will pick the ball. Most people would strictly prefer urn I. If they were 'probability-sophisticate', this would imply that they think that there are fewer than 50 red balls in urn II. They should therefore rather prefer to bet on drawing a black ball from urn II. This paradox suggests that people are not probability sophisticated and dislike a situation where probabilities are unknown, i.e. they are ambiguity-averse.

The observation that participants in choice-experiments tended to deviate systematically from the prediction of expected utility theory has led to the hunt for a descriptive theory of choice under risk and uncertainty accommodating non-linear probability weighting, heterogeneous valuation of bad and good events, and ambiguity aversion. The prospect theory proposed by Kahneman and Tversky (1979) gained a wide acceptance in the field of non-expected utility theory and could provide a worthy alternative to expected utility theory. Nevertheless, despite the large set of experimental evidence showing that alternate models are better for choice evaluation under conditions of risk and uncertainty, the use of non-expected utility models in empirical studies is still rare in agricultural and development economics. We will therefore not delve into more details. Interested readers are referred to literature reviews from Etner et al. (2012) and Starmer (2000) and the books from Gilboa (2009) and Wakker (2010) for a in depth treatment of this subject matter. We turn now to estimation of risk exposure with empirical data.

1.3 Estimating risk exposure

The goal of risk estimation in development and agricultural economics could be summarized as the estimation of the mean, variance and skewness of the probability distribution of production, yield, profit or consumption. More generally, the mean, variance and skewness are defined as the first, second and third central moments of a probability distribution. We will use both terminologies below. The stochastic production approach has been the dominant approach for estimating risk in the agricultural economics literature. The key insight of Just and Pope (1978) was to split the production function into a deterministic part and a stochastic part, allowing inputs to be risk increasing, risk neutral or risk decreasing:

$$y = f(x, \beta_1) + h(x, \beta_2)\varepsilon \tag{5}$$

where y is the quantity produced, $f(x, \beta)$ is the mean, x the input and $h(x, \beta)^2\varepsilon^2$ is the variance. As Antle (1983) put it, the idea is basically to specify a deterministic model to which an error term is appended. Antle (1983) showed that although input effects on variance are not determined by their effects on the mean, this specification restricts the inputs' effect to have the same direction on variance and higher moments.

Antle (1983) proposed an alternative model wherein the central probability moments are directly specified as:

$$\mu_1(x, \beta_1) = \int y f(y|x) dy \tag{6}$$

$$\mu_i(x, \beta_i) = \int (y - \mu_1)^i f(y|x) dy \qquad i \geq 2 \tag{7}$$

where β_i relates the input x to the moment μ_i. This approach relaxes the cross-moment restrictions: the inputs' elasticity with respect to variance does not restrict their elasticity with respect to higher moments[2]. Antle (1987) showed that the different central moments can be estimated using feasible Generalised Least Squares estimators.

Antle (2010) extended his model to allow inputs to have asymmetric effects on the central moments of the probability distribution. The approach is based on the estimation of *partial* moments, i.e. probability moments below and above a given threshold of the yield or output. Once the effect of a given input on the higher partial moments is netted out from its effect on the lower partial moments, the conclusion regarding the risk effects of this input can change radically. It is illustrated with data from Ecuador where labour is shown to be risk increasing according to the central moments approach and to be risk decreasing according to the partial moments approach.

Kim et al. (2014) recently proposed an estimation framework based on quantile regressions. The first stage consists of estimating a production function. The second stage consists of using the residuals of this first stage regression as dependent variable in a series of quantile regressions. The authors also provide a derivation of the risk premium in terms of the quantiles of the probability distribution. Based on the analysis of farm data in South Korea, they show that 90% of the cost of risk, defined as the risk premium, comes from downside risk. The method has also been applied to compare the effect of genetically modified crops on corn yield distribution (Chavas and Shi 2015).

The models presented above rely on the residual of the production function to estimate risk. By contrast, in the field of efficiency analysis of agriculture production, a different interpretation is given to the residual: it represents farmers' inefficiency. The field of stochastic frontier analysis sought to

[2] The model from Antle (1983) has been labelled the *flexible method of moments* because it relaxes the assumption on cross-moment elasticity of inputs, and more recently, the *linear method of moments* because of its assumption that the central moments of the probability function can be approximated by a linear function (Tack et al., 2012).

unify both approaches by extending the stochastic production function from Just and Pope with a one sided error terms capturing the distance from the efficiency frontier (Battese et al. 1997; Sabul C Kumbhakar 1993; Subal C. Kumbhakar 2002; Subal C. Kumbhakar and Tveterås 2003; S. Kumbhakar and Tsionas 2010) or by estimating risk as the variance of the inefficiency terms (Bera and Sharma 1999; Jaenicke et al. 2003; Wang and Schmidt 2002). As the residual is ultimately determined by what is left out of the regression model, the decomposition of the residual into an inefficiency term, risk term, misspecification error and noise is made difficult: 'whether the residuals represent the ignorance of the firm under study or [the ignorance of] the analyst' is often unclear (Saastamoinen 2013).[3] This is likely one of the major weaknesses of risk estimation models based on the error term.

A novel semi-parametric approach has been proposed by Tack et al. (2012) to circumvent this issue. Instead of estimating equation (6) and (7), Tack et al. (2012) estimate the raw moments of the conditional distribution of output via the following equation:

$$y_{it}{}^{j} = f(x, \beta_1) + \varepsilon_{it} \tag{8}$$

where y_{it} is the output at time t for observation i put to the power j and each j equation corresponds to each j raw moment. These raw moments provide all the required information to estimate the mean, variance and skewness of the conditional distribution, estimated as a maximum entropy density distribution. This distribution contains as a special case the normal, beta, chi-square, beta and many other distributions. The advantage of the raw moments approach is to decrease the risk of obtaining biased results caused by misspecification of the first moment equation, i.e. equation (5) in the model of Just and Pope and equation (6) in the model from Antle (1983). This method has notably been used to estimate the effect of the El Niño Southern Oscillation on US corn production and downside risk (J. B. Tack and Ubilava 2013).

The availability of climate data has led to an explosion of the number of studies on the link between climate and the economy. This 'new climate-economy literature' (Dell et al. 2014) does not seek to recover conditional output distribution and is hence not focused on risk *per se*, but rather on how weather shocks affect various dimensions of the economy, from aggregate output, labour productivity, health and mortality, conflict and political stability, and, naturally, agriculture.

In order to estimate the relationship between weather shocks and the economy, these studies use a similar reduced-form panel specification which can be generalized as (Dell et al. 2014):

[3] Interested readers are referred to the literature review of Saastamoinen (2013) for a detailed presentation of the issue.

$$y_{it} = \beta \boldsymbol{C_{it}} + \gamma \boldsymbol{Z_{it}} + \mu_i + \theta_{rt} + \varepsilon_{it} \quad (9)$$

where y_{it} is the outcome of interest for unit i at time t, $\boldsymbol{C_{it}}$ is the climate variable, $\boldsymbol{Z_{it}}$ is a set of control variables, μ_i is a unit fixed effects, θ_{rt} is a time fixed effects which might vary between regions, r, and ε_{it} is an error term. Most of these studies, when focusing on agriculture, are conducted at the aggregate level (county, district, state or country level) with the few exceptions such as the studies of Yang and Choi (2007) and Welch et al. (2010) where farm and household level data are used. The outcome variable is either yield, profit or revenue and the climate variable is either temperature, precipitation or a combination of both.

As the climate variable is exogenous and varies randomly over time, the studies based on equation (9) do not suffer from reverse causality bias. Furthermore, the use of fixed effects, μ_i, and time fixed effects, r, account for possible omitted variable bias (e.g. unobserved spatial characteristics) and allow for different trends across subsamples. Studies based on equation (9) 'makes relatively few identification assumptions and allows unusually strong causative interpretation' (Dell et al. 2014). Dell et al. (2014) provides an extended survey of the new climate-economy literature.

Most of these studies use a non-linear specification of the climate variable, $\boldsymbol{C_{it}}$, be it temperature or precipitation, with a preference for quadratic specification: precipitation or temperature increases the output variable, y_{it}, up to a point where higher precipitation or temperature has an adverse impact on y_{it}, each additional degree or precipitation amount decreasing y_{it} (e.g. Hidalgo et al. 2010; Lobell et al. 2011a; Lobell et al. 2011b; Schlenker and Roberts 2009; Schlenker and Lobell 2010). Another interesting insight from the 'new climate-economy' literature is the clear distinction between weather, defined as the realisation of the distribution of possible weather events, and climate, defined as the distribution of these possible weather events.

We reviewed above some of the core techniques for estimating risk at the farm level. Most models are based on the residual of a regression of profit or output, with the exception of the raw moments approach of Tack et al (2012). Recently, the availability of climate data has led to the burgeoning studies on the link between various aspect of the economy and weather shocks. The latter studies are however not concerned *per se* with risk; rather, they look at the impact of shocks. The contribution of chapter 2 is to propose a model based on climate data and a model similar to equation (9) in order to assess climate risk exposure and vulnerability.

1.4 Empirical studies on the link between risk and poverty

The absence of effective risk management strategies leaves farmers exposed to the vagaries of the weather, resulting in large fluctuations in agricultural production, income and household consumption. In Zimbabwe, for instance, Kinsey et al. (1998) report that household maize production dropped from more than 3 ton in 1991 to half a ton during the 1992 drought. In 1981-1985 in Burkina Faso, a period marked also by a major drought, the standard deviation in crop income was half the long term average income (Kazianga and Udry 2006). Furthermore, half of the variation in crop income was directly passed into consumption, resulting in median caloric intake 30% below the World Health Organization's recommendations. In Ethiopia, even relatively common weather conditions, such as a rainfall deficit expected to occur every five years on average, can decrease consumption by 10 to 20 percent (Porter 2012). The impact of a drought is rarely felt equally by all members of the household. Hoddinot (2006) shows that during the 1994-1995 Zimbabwean drought, women and children under the age of 12-24 months bore the brunt of the shock, the males' body mass index staying relatively stable over the same period. As there are large seasonal variations in rural households' consumption and, as data collection occurs only once a year at best, a large part of consumption fluctuations might even go unnoticed in most studies (Dercon and Krishnan 2000).

Weather shocks not only bring short term fluctuation in income and consumption, they can also have long lasting consequences on households' well-being by pushing them in poverty traps, i.e. 'equilibrium levels of poverty in which one may slide relatively easily, but from which there is no possible recovery without *outside* intervention' (Dercon 2008). It took Ethiopian households on average 10 years to rebuild their cattle herd following the 1983-1985 drought, while consumption growth in the 1990's of the most affected households remained 16% lower than the least affected (Dercon 2004). The 1984-1985 Ethiopian famine had a large and irreversible impact on children under the age of 36 months at that time. Dercon and Porter (2014) observe that adult height shortfall caused by poor infant nutrition is associated with a 3-8% p.a. income loss over adult life. Even marginal changes in rainfall have long-lasting effects. Although the 1994-1995 Zimbabwean drought was a relatively mild drought by African standards (no famine or emergency appeal for food aid), it lowered the growth velocity of children aged 12-24 months at that time (Hoddinott 2006). While the children living in wealthier households were able to recover, those living in poor households never did.

The 2002 drought in Ethiopia is considered to have been well-managed: no famine death was reported despite it being of a similar magnitude than the 1983-1985 one. Two years later, the 2002 drought nevertheless still had a negative and significant impact on consumption (Dercon et al. 2005). More generally, a 10 percent decrease in rainfall depresses consumption growth by 1 percentage point four to five years later in Ethiopia (Dercon and Porter 2014). This high level of path dependence in poverty has been observed in several other studies (e.g. Bigsten and Shimeles 2008; Yesuf and Bluffstone 2009).

Households do not stay passive: they implement a wide array of strategies to deal with these shocks and decrease risk exposure. These strategies are generally classified into two types: *ex-ante* risk management strategies and *ex-post* risk coping strategies (Alderman and Paxson 1992). The aim of *ex-ante* strategies is to decrease income risk (e.g. decrease variance and skewness of the conditional distribution of income). The aim of *ex-post* strategies is to reduce the impact of income shocks on consumption. *Ex-ante* strategies include the diversification of income sources and the adoption of low risk production techniques. *Ex-post* strategies include self-insurance in the form of asset holding ready to be sold in case of adverse shocks and informal insurance via solidarity networks of friends and family ready to be called upon in case of need.

Theses ex-post and ex-ante strategies can have several adverse consequences such as holding a sub-optimal amount of productive assets (e.g. cattle) or the selection of activities less sensitive to rainfall variations but less profitable, contributing to the trapping of households in poverty. Although *ex-ante* and *ex-post* strategies should not be analysed separately as the availability of *ex-post* strategies typically influence the choice of *ex-ante* strategies (e.g. Dercon 1996; Rosenzweig and Binswanger 1993), we discuss each type of strategy separately for the sake of clarity. We start by surveying the literature on risk aversion as risk aversion determines the price one is ready to pay to decrease risk.

1.4.1 Risk aversion

There are two main approaches for estimating risk aversion: analysing production choices or conducting choice-experiments. Models using production data are based on the insight that risk-averse farmers will choose a mix of inputs that decrease risk at the cost of forgone output, whereas risk-loving farmers will choose a riskier input mix with higher returns. Two sets of parameters need hence to be estimated: the technology parameters which describe the effect of inputs on risk and profit and the utility parameters mapping the input mix into risk aversion. These two stages can be carried either recursively (John M. Antle 1987, 2010; Foudi and Erdlenbruch 2011; Groom et al. 2008; Simtowe et al. 2006) or jointly (e.g. Chavas and Holt 1996; Koundouri et al. 2009; Subal C. Kumbhakar 2002; Subal C. Kumbhakar and Tveterås 2003; S. Kumbhakar and Tsionas 2010; Love and Buccola 1991; Pope and Just 1991; Saha et al. 1994).

The core assumption of these models is that producers maximise expected utility. Most of these studies, carried out in developed countries, points toward a DARA and IRRA risk preference structures.

A large heterogeneity in risk preferences has been found among farmers' populations. Antle (1987) finds that farmers in the Aurepalle village in India were both Arrow-Pratt and downside-risk-averse but at very different levels, ranging from risk neutrality to risk premiums as high as 25%. Lence (2009) calls into question the validity of estimating risk aversion with production data, particularly when shocks are not large or the sample is small, which is the case in most of the studies cited above. Furthermore, Just et al. (2010) show that although the models can be locally identified, they are not globally identified: an infinite set of pairs of technology and utility functions can equally well fit the data. These two articles are a serious blow to the whole field of risk preferences estimation based on production data.

The use of choice-experiments to recover risk aversion in developing countries dates back at least to the 1980 Binswanger study on risk aversion in India (1980). The use of choice-experiment has provided a fertile ground for testing the validity of the expected utility theory in developing countries.

Deviation from expected utility theory has been found in Uganda by Humphrey and Verschoor (2004) and by Mosley and Vershoor (2005) in Ethiopia, Uganda and Andhra Pradesh (India): rank dependent utility function and non-linear probability weighting function were found to explain better the data. Harrison et al (2010), with the same data than Mosley and Vershoor (2005), show by contrast that there is equal support both for expected utility theory and prospect theory. When both models are used simultaneously to explain the choices made in the experiment, they find that subjects behaving according to expected utility exhibit risk aversion and subjects behaving according to prospect theory exhibit risk loving behaviour. Risk aversion has also been found to vary geographically. Tanaka and Munro (2014), with results holding both under expected utility theory and prospect theory, found that farmers in the less favourable agro-climatic zones in Uganda were also the most risk averse. Several authors also find that the exposure to large shocks such as a severe drought was correlated with higher risk aversion (Gloede et al. 2013; Yesuf and Bluffstone 2009). Choice experiments are however not immune to bias. Cilliers et al. (2015) showed for instance that in an experiment in Sierra Leone the presence of a foreigner, typically a white researcher from a Western country, altered the behaviour of participants.

Mosley and Verschoor (2005) suggest that there could be a mutually self-reinforcing cycle between risk aversion and poverty. Ex-ante risk management strategies by asset-poor households are not adequate to deal with important shocks. Asset-poor households have hence no other choice than to deplete their asset stock in order to smooth out the effect of shocks. This increases the risk of failing into chronic poverty, which, in turn, increases risk aversion, reducing their willingness to undertake riskier and more profitable investments or to adopt new technology[4]. As a result, this further traps them in poverty. Dercon (2008) stresses rather the lack of insurance, credit and investment options among the poor as the main explanation behind the uptake of costly ex-ante risk management strategies. Innate differences in risk aversion between the rich and the poor would not play a central role in the larger adoption by the poor of costly *ex-ante* risk management strategies.

1.4.2 Ex-ante strategies

As discussed in section 1.2, risk might have a detrimental effect on well-being as it is hard to reduce risk without reducing expected income. This results in 'efficiency losses, more inequality, [..] and poverty traps.' (Dercon 2002). We discuss below one of the main ex-ante strategies: diversifying the source of income. The choice of low-risk low-return income generating activities constitutes a second major ex-ante strategy. We will discuss it in section 1.5 when presenting the role of risk aversion in the low rate of technology adoption.

Diversification of the household activities into uncorrelated income sources is a key strategy to reduce risk. Non-farm income may already account for more than 40 percent of average household income (Barrett et al. 2001). However, risk mitigation cannot explain all diversification as 'it is widely believed that risk aversion decreases with wealth but it is observed that diversification increases with wealth' (Barrett et al. 2001).[5] One reason for the lower diversification among the poorest households is the barriers to entry into non-farm enterprise such as capital constraints and education (Reardon et al. 2000). Furthermore, finding income sources not correlated with the agricultural sector is difficult in rural areas. Fafchamps et al. (Fafchamps et al. 1998) find that 'droughts adversely affect not only crop income but also non-farm income'. The authors state that these results are 'consistent with Sen (1981) who remarked that droughts lead to a collapse of the demand for local services and crafts'.

Several authors have shown that preserving biodiversity could be an effective ex-ante risk management strategy, as ecologically diverse agricultural systems tend to be both more resilient to climate shocks, pest invasion and crop diseases and more productive (Di Falco and Perrings 2005; Di Falco and Chavas 2006, 2009; Smale et al. 1998). As bio-diversity plays a role in ex-ante risk mitigation, risk aversion has been found to increases bio-diversity both in developing and developed countries (Bezabih and Sarr 2012; Di Falco and Perrings 2005; Nastis et al. 2013).

[4] The impact of risk aversion on technology adoption is discussed in more detail in section 1.4.3.
[5] Several other factors explain diversification (Barrett et al. 2001). The *push factors* include risk diversification, ex-post coping strategies such as looking for alternative income or food consumption source after a crop failure, limited access to inputs forcing households to diversify their activities. The *pull factors* include complementarities between activities (e.g. intercropping), economy of scope or skill endowments.

Before presenting the ex-post risk coping strategies, let us consider for a moment the challenge of preserving the environment in developed countries. Many OECD countries have set up agri-environmental schemes (AES) aimed at financially rewarding farmers for adopting eco-friendly practices. It is however as yet unclear what impact they have on risk exposure. Some production standards imposed by AES, such as a reduction in fertiliser and pesticide application rates, might have a direct impact on risk exposure. Indeed, organic and low input farmers are generally more exposed to production risk than conventional farmers (Berentsen et al. 2012; Finger 2014; Gardebroek 2006; Serra et al. 2008). Corroborating these findings, organic farmers tend to be less risk averse than conventional farmers (Gardebroek 2006; Serra et al. 2008). Morris et al. (2000) report a concern among English farmers that the rigidity of their local AES reduced their ability to take remedial action in case of pest infestation or severe weed events. Similarly, AES contract length was found to affect negatively the decision to join an AES because it tied up farmers' hands over a long period of time (Peerlings and Polman 2009). The change in agricultural practices required to join an AES is also generally perceived as a risk (Wynn et al. 2001). It is however as yet unclear if joining an AES is objectively riskier or if it is perceived as such because of the uncertainty linked to the change in long tried farming practices. We will investigate this in more detail in chapter 5, studying the effect of joining an AES on Irish farmers' risk exposure.

1.4.3 Ex-post strategies

A typical feature of weather shocks is that they tend to be covariate: many households are hit at the same time. While households manage to deal relatively well with idiosyncratic shock, i.e. shocks affecting only one household at a time (e.g. illness) (e.g. Porter 2012), they are ill-equipped to deal with covariate shocks such as drought or price swings. Naturally, successive shocks are more difficult to deal with than a single one (Alderman 1996, cit. in Dercon, 2002). We review below two important *ex-post* strategies adopted by households to deal with shocks: self-insurance in the form of asset holding, risk sharing via solidarity networks of friends and family. We then briefly comment the link between ex-post coping strategies and environmental degradation and a weather index insurance.

One the main strategies for coping with adverse shocks is asset holding, typically in the form of cattle, crop inventory or jewellery. In case of need, households can sell an asset in order to buffer the effect of the shock. It is hence often described as a self-insurance. However, their effectiveness at decreasing the impact of covariate shocks is limited, as shown in the case of the 1981-1985 drought in Burkina Faso where livestock sales compensated at most 30% of village-level income shocks (Fafchamps et al. 1998). Running down crop inventory or decreasing food consumption was more a common response than selling livestock (according to the study from Kazianga and Udry, 2006, with the same dataset).

When all households seek to sell livestock at the same time, as when they are all affected by large drought, the terms of trade for livestock against food will collapse.

The *lumpiness* of asset limits also their use for consumption smoothing: they are very costly to acquire, representing often several months of crop income, and are hence not easily disposed-of. Furthermore, livestock are key productive inputs used to plough the fields or providing manure for the crops. Hoddinott (2006) finds that in Zimbabwe, during the 1994-1995 minor drought, more than half of the households owning at least two oxen sold at least one ox. By contrast, only 15% did so when they owned only one or two oxen. Disposing of all the oxen can indeed have long-lasting consequences such as forcing households to rent each following year an ox at a high price, decreasing their ability to save money in order to buy another ox and get back on their feet. The risk of falling into such asset poverty traps could explain why asset rich households pursue consumption smoothing while asset poor households pursue asset smoothing (e.g. Zimmerman and Carter 2003).

However, some behaviour observed during time of famine is hard to reconcile with the idea that asset-poor households maximise the expected utility of consumption over time by protecting their assets. During the Ethiopian famine of 1984-1985 for instance, many households declined to sell their livestock even when it could have saved their family members, and their own, lives. Dercon (2008) suggest that non-expected utility models of choice under risk and uncertainty could explain this behaviour: individuals would exhibit risk-loving attitudes when faced with large shocks, clinging to the hope of conserving their *status quo*, however remote is this possibility.

A second important *ex-post* strategy is to rely on the solidarity of friends and family as a form of informal insurance. Its effectiveness is however limited mostly to idiosyncratic risk. Bramouillé and Kranton (2007) show that social networks provide higher welfare gains when they link individuals across different villages and communities, but geographical proximity and kinship are the major determinants of risk-sharing network formation (Fafchamps and Gubert 2007). The result is that risk-sharing networks rarely include people with uncorrelated risk (Fafchamps and Lund 2003). They are hence of limited use against covariate shocks such as drought or price swings. As all the individuals in the social network are affected at the same time, there is no surplus to share. Several studies tested and rejected the hypothesis of full risk pooling at the village level (e.g. Townsend 1994). Kazianga and Udry (2006) find almost no risk sharing in Burkina Faso over the 1981-1985 period. Reardon et al. (1988, cited in Dercon, 2002) report that transfers represented 'only 3 per cent of the losses for the poorest households in the Sahel' during the 1981-1985 drought. Nor does Yilma (2014) find evidence for reliance on gifting from friends and family to deal with shocks in Ethiopia, in a survey conducted in

2011. Morduch (1999) gives many other examples of a limited role for transfers in smoothing income shocks.

A disadvantage of these risk sharing networks is that they can lead to unequal patron-client relationships between the poorest households and wealthier ones (Fafchamps 1992) when the poorest households are not simply excluded (e.g. Santos and Barrett 2011). Furthermore, Di Falco and Bulte (2011) find that forced sharing norms in social networks diverted investment away from sharable liquid assets toward assets not sharable, but with lower returns. Jakiela and Ozier (2016) find that sharing norms 'distort incentives towards less visible, but potentially less profitable, investments, and may consequently slow economic growth'. Baland et al. (2011) provide anecdotal evidences that some individuals in Cameroon attempt to fend off network requests by contracting costly credit in order to 'pretend to be poor'. As Chuang and Schechter (2015) put it, 'this suggests that we should be interested both in how social networks work as a conduit for financial transactions, but also how social networks enforce these transactions'. There could be indeed a 'dark side' to social capital (Di Falco and Bulte 2011).

Lastly, households may rely on income or food sources with detrimental effects on the environment such as wildlife poaching (Barrett and Arcese 1998), firewood and charcoal selling (Little et al. 2001) The resulting environmental degradation can then itself increases risk exposure (e.g. decreases in water availability due to deforestation) and contribute to establish a mutually self-reinforcing cycle between poverty and environmental degradation (Barbier 2010; Barrett et al. 2011; Dasgupta et al. 2005).

A new blend of insurances has emerged on the policy arena since the beginning of the nineties: weather index insurance and affiliated instruments, also known under the generic term of *index-based risk transfer product* (IBRTP). The concept of index insurance is that the indemnities are not paid according to the actual loss but on the basis of an index highly correlated with the loss. Examples of such indexes are rain gauges, area-yield, wind speed or bio-mass as captured by satellite imagery.

Index insurance might prove adequate in developing countries where governments can ill-afford the heavy subsidies needed for standard multiple peril crop insurance (Ibarra and Skees 2007). Weather index insurance is indeed much cheaper. For instance, rainfall insurance could be triggered objectively and remotely monitored based on satellite or rain gauge data. There is no need to send an employee to check that each policy holder is affected by a drought. Furthermore, it decreases the cost linked to asymmetric information (moral hazard and adverse selection) as the risk can be objectively assessed thanks to historical rainfall data. However, the uptake of weather insurance has been lower than expected, the explanations ranging from a lack of information, the complexity of the product or behavioural biases.

Clark (2011) showed recently that the basis risk was the main determinant of non-adoption. Hill et al. (2013) find that wealthier households were more likely to purchase weather index insurance and that risk aversion decreased the propensity to adopt it. Dercon et al. (2014) found that basis risk in weather insurance makes it a complement to informal risk sharing, implying that weather insurance could reinforce existing informal risk sharing networks and that communities with well-functioning informal risk sharing networks should be targeted first for weather insurance. Despite these promises, weather index insurance availability is still very low in Sub-Saharan Africa. As reviewed above, the traditional *ex-post* strategies adopted by households are not very effective at dealing with covariate and recurrent shocks such as drought and price swings.

To sum up, traditional on-farm risk management strategies might be effective against idiosyncratic risk such as illness, but they might prove very costly, lock farmers in poverty traps, increase inequality and are not adapted to covariate shocks such as drought. Lastly, recent studies point toward a positive relationship between bio-diversity preservation and risk management and the development of weather index insurance could provide an effective way for farmers to deal with risk.

1.5 Adoption of agricultural technologies and their impact on well-being

New agricultural technologies, such as drought tolerant crop varieties, could play a crucial role in improving food security in Africa. High yield varieties and the adoption of modern farming practices drove the Green Revolution in Asia and could provide increases in agricultural productivity across Africa as well. The uptake of new agricultural technology in Africa is however still limited and far from complete (Foster and Rosenzweig 2010).

One key reason for adopting a new technology is its profitability. The return on adoption is however not easy to estimate (Foster and Rosenzweig 2010). Even in the relatively simple case of profit maximising entities, where the return on adoption could be measured as the increase in profit caused by the new technology, some inputs might hard to measure, particularly when considering small-scale farmers in developing countries where a large part of the input invested on the farm consists of home-labour, which is hard to measure and hard to value. For instance, Foster and Rosenzweig (2010) notes that Duflo's study (2008) of the impact of fertiliser adoption on output neglects changes in labour provision. As the change in labour provision represents a cost, the benefit from the adoption might be overestimated if not properly taken into account.

There is also a large heterogeneity between farmers in terms of plot quality, agro-ecological conditions, management skills and access to capital and inputs. These differences might be hard to assess. The observation that high yield farmers use improved seeds does not imply that farmers with traditional seeds should use improved seeds. It might be the case that farmers with improved seeds would also perform better with traditional seeds (heterogeneity between farmers) or that traditional seeds are better suited to the marginal lands used by farmers with low yields (heterogeneity in the return on adoption). An estimation of the impact of a new technology based on observational data which does not properly account for the heterogeneity between adopters and non-adopters will be biased: a part of the estimated increase in yield attributed to the new technology is, in fact, caused by the better management skills of the adopters or the higher suitability of their plots, for instance.

These issues have been dealt with via endogenous switching regressions (Asfaw et al. 2012a), propensity score matching methods (Becerril and Abdulai 2010; Kassie et al. 2014; Mendola 2007), the use of both methods side by side (Amare et al. 2012; Asfaw et al. 2012b; Khonje et al. 2015; Shiferaw et al. 2014), or panel data analysis (Bezu et al. 2014; Mathenge et al. 2014; Smale and Mason 2014; Suri 2011). In endogenous switching regression methods, the decision to adopt and the impact of adoption on output are analysed jointly. The identification of the causal effect of adoption on the yield rests on the assumption that at least one variable used to explain adoption is not correlated with yield (exclusion restriction). In the propensity score matching methods, a set of observable variables is used to build a comparable set of pairs of adopters and non-adopters. The assumption is that conditional on these variables, the decision to adopt is random. However, some characteristics determining adoption might be unobservable to the analyst, implying that the conditional independence assumption is violated.

Most studies find a positive effect of improved seeds on farmers well-being, be it in terms of income, food security or reduced depth of poverty. For instance, Minten and Barrett (Minten and Barrett 2008) find that areas that ‘have higher rates of adoption of improved agricultural technologies and, consequently, higher crop yields enjoy lower food prices, higher real wages for unskilled workers, and better welfare indicators’. Some studies find also that the poor benefit particularly (Bezu et al. 2014; Mathenge et al. 2014). However, gains might vary across farmers, depending on geophysical conditions, farm size, and other characteristics affecting output. Heterogeneous treatment effects associated with adoption have been analysed in detail for the case of hybrid maize in Kenya by Suri

(2011). Suri shows that the return on adoption might be null for a large share of the Kenyan farming population.

The studies have provided a plethora of explanations for the low adoption of modern agricultural technologies. Following a literature review from 1985 (Feder et al. 1985), they include: farm size, risk exposure and risk aversion, human capital, capital and credit constraints, limited access to information. Studies on the determinant of the adoption of modern seeds performed since 2000 show that farmers still face similar constraints: the high price of seeds (e.g. Fisher et al. 2015; Wekesa et al. 2003), their low availability (e.g. Amare et al. 2012; Asfaw et al. 2012a; Fisher et al. 2015; Wekesa et al. 2003), a lack of capital and credit access (e.g. Fisher et al. 2015; Lambrecht et al. 2014; Zeller et al. 1998), inadequate information about their usage and yield (e.g. Amare et al. 2012; Asfaw et al. 2012a; Fisher et al. 2015). Other studies have shown that the low rate of adoption could be explained by behavioural bias such as high rate of time discounting (Duflo et al. 2008, 2011) while others still have shown that the improved seeds were not profitable for a large share of the population so that farmers may behave perfectly rationally when deciding not to adopt (Suri 2011).

Risk aversion and risk exposure has also been shown to play an important role in the adoption decision (e.g. Richard E Just and Zilberman 1983). New technology such as fertilisers or improved seeds represent a significant investment for poor farmers, the benefits of which might be hard to ascertain without prior experience. Simtowe et al. (2006) showed that a high level of risk aversion, among other factors such as education, leads to a low level of adoption of hybrid maize among farmers in Malawi while Knight et al. found similar results in Ethiopia (J. Knight et al. 2003). Hill (2009) finds that more risk averse farmers were less likely to invest in a profitable but risky crop in Uganda (coffee), a fact which was particularly salient for poor farmers unable to insure against shocks.

Brick and Visser (2015) also found that farmer with a high degree of risk aversion were more likely to use traditional agriculture techniques and less likely to use modern seeds even when insurance is available. By contrast, Dercon and Christiaensen (2011) and Berhane et al. (2015) find that fertiliser use in Ethiopia increases if insurance is offered. Karlan et al. (2014) even go a step further by showing that in Ghana the binding constraint for investing in new technology is risk and not credit access: when relaxing risk constraint by providing weather index insurance, households were actually able to find the credit required to increase expenditure on farm inputs. Zeller et al. (1998) suggest that these risk-induced poverty traps are not inescapable, as high risk averse farmers in Malawi were able to adopt new technology, such as hybrid maize and tobacco, provided that 'policies improve their access to credit, extension, input and output markets'.

Education has also been shown to play a major role in the adoption process as it facilitates learning about the use and potential of new technologies. Weir and Knight (2000) show that education had two positive externality effects: 'educated farmers are early innovators, providing an example which may be copied by less educated farmers; and educated farmers are better able to copy those who innovate first, enhancing diffusion of the new technology more widely within the site'. In related study, Knight et al. (2003) found that education reduced risk aversion and hence increased adoption.

It is has also been long recognized that social networks play an important role in the diffusion of technology. For instance, Besley and Case (1993), in their discussion of the modelling and estimation of farmers' adoption decisions, stress the importance of learning from the experience of early adopters and from discussion with peers. Several empirical studies confirm the importance of the learning effects, mediated by peers and extension agents, on the decision to adopt (e.g. Bandiera and Rasul 2006; Conley and Udry 2010; Isham 2002; Moser and Barrett 2006; Munshi 2004).

Furthermore, Foster and Rosenzweig (1995) found that the profitability of a seed increased with neighbours' experience because better knowledge on the management of the new seed was gained by observing peers. More recently, Van den Broeck and Dercon (2011) found similar effects among Banana growers in Tanzania and showed that the impact of information depended on the network type, the kinship network being the most effective at transforming information into higher yield. Barr (2000) also found some productive effects of social capital in the Ghanaian manufacturing sector thanks to better flows of technological information. Nevertheless, the impact of social networks on the adoption process might be complex. Bandiera and Rasul (2006) find that the probability of adoption follows an inverse-U shape in the size of the social network: when there is only a few adopters in the network, knowing one more adopter increases the probability of adoption; when there are many adopters in the network, some farmers might have an incentive to free-ride and rely on the solidarity of their peers.

We propose in chapter 3 an impact evaluation framework based on a set of randomized controlled trials. It offers unbiased estimate of the effect of improved seeds on yield and allows us to estimate the role played by the input adjustment of farmers in yielding the full potential of the improved seeds.

1.6 Conclusion

Farmers in Sub-Saharan Africa are largely exposed to the vagaries of the weather. In order to deal with this risk, they adopt various *ex-ante* risk management and *ex-post* risk coping strategies. However, *ex-post* strategies tend to break down when large covariate shocks, such as droughts, occur; while *ex-*

ante strategies tend to be very costly and might contribute to trap households in poverty. Improved seeds and fertilisers could go a long way in improving food security. However, they are not widely adopted. A myriad of factors explain this situation, but risk and the heterogeneity of new technology returns likely play an important role. Although solidarity networks of kin and friends are generally inadequate to deal with large covariate shocks, they are a key channel for technology diffusion. Nevertheless, an emerging body of literature shows that the social pressure to share among social networks might also distort incentives to invest in lucrative activities because of the fear of being taxed by kin or friends. Lastly, while bio-diversity has been found to decrease risk exposure in developing countries, risk aversion might by contrast deter some farmers from adopting eco-friendly practices in developed countries. We summarize below the contribution of the current thesis to the various fields of research covered in this literature review.

Chapter 2 aims at providing a simple framework to assess the climate risk exposure by combining climate data and household level consumption data. Instead of relying on the residual as in the classic risk estimation framework of the agricultural economic literature (e.g. John M Antle 1983; john M. Antle 2010; Richard E. Just and Pope 1978; Kim et al. 2014), we propose to estimate the impact of weather shocks on consumption variables with a regression model similar to the one used in the 'new climate economy' literature (Dell et al. 2014). The novelty comes from the use of the statistical properties of the Standardized Precipitation Evapotranspiration Index (SPEI), our climate variable, in order to estimate the shape of the conditional distribution of consumption and provide direct estimate of its variance, its skewness. Poverty risk and other vulnerability indices could be computed with this approach.

Chapter 3 provides an analysis of the contested issue of the profitability of improved seeds. It is likely that no single factor can, on its own, explain the puzzle of their low adoption in Sub-Saharan Africa. However, Suri (2011) showed that some agricultural technologies might not be profitable for a large part of the population, while Rosenzweig and Foster (2010) highlight the fact that many studies may overestimate the benefit of adoption because they neglect farmers' adjustment cost, i.e. they measure the total effect of adoption instead of the net effect. Based on open and double-blind randomized controlled trials (RCT) in Tanzania, we provide an assessment of the role of farmers' behavioural responses in driving the increase in yield of improved seeds. We find that at least 50% of the increase in yield estimated in a traditional RCT would not materialise without an increase in labour, land allocation and other dimensions of effort. Our experimental design provides also a lower and higher bounds of the net effect improved seeds.

Based on the data from the RCT cited above, chapter 4 investigates if the participants adopt an evasive behaviour to fend off solidarity requests from their social network when exposed to a positive income shock, i.e. receiving improved seeds. We use here only the open RCT conducted in Tanzania, not the double-blind one: we want to observe the participants' behaviour when they know that they have a positive income shock. The evasive behaviour is measured among farmers having received the improved seeds as a decrease in social interactions which could reveal to their social network that they received the improved seeds. We find that the propensity to adopt an evasive behaviour increases with the size of the kin network. To our knowledge, this constitutes the first set of evidence, based on a RCT design involving real interactions - i.e. not in a choice experiment as in Jakiela and Ozier (2016) or with observational data as in Baland et al. (2011) and Di Falco and Bulte (2011) - that hiding from the social network tax takes place in a village economy. Nevertheless, we stop short of concluding that this evasive behaviour bears any economic consequences. Indeed, although they asked for less help on the experimental plot at harvest time, the participants did not ask for less help at planting and weeding time where a decrease in labour could have decreased yields.

Lastly, chapter 5 investigates the impact on Irish farmers' risk exposure of the Rural Environment Protection Scheme (REPS), an agri-environmental scheme. Organic and, more generally, low input agriculture tends to increase risk exposure according to several studies (e.g. Berentsen et al. 2012; Finger 2014; Gardebroek 2006; Serra et al. 2008) while risk aversion is believed to play a role in the low adoption of sustainable production techniques (e.g.Gardebroek 2006; Morris et al. 2000; Peerlings and Polman 2009). We show that REPS does not increase risk exposure, and adequately compensates farmers for foregone returns. Although we do not analyse the decision to join REPS, this could be one of the reasons of its large success.

1.7 References

Alderman, Harold (1996), 'Saving and economic shocks in rural Pakistan', *Journal of Development economics,* 51 (2), 343-65.

Alderman, Harold and Paxson, Christina H (1992), *Do the poor insure?A synthesis of the literature on risk and consumption in developing countries* (164: World Bank Publications).

Allais, M. (1953), 'Le Comportement de l'Homme Rationnel devant le Risque: Critique des Postulats et Axiomes de l'Ecole Americaine', *Econometrica,* 21 (4), 503-46.

Amare, Mulubrhan, Asfaw, Solomon, and Shiferaw, Bekele (2012), 'Welfare impacts of maize–pigeonpea intensification in Tanzania', *Agricultural Economics,* 43 (1), 27-43.

Antle, John M (1983), 'Testing the stochastic structure of production: a flexible moment-based approach', *Journal of Business & Economic Statistics,* 1 (3), 192-201.

Antle, John M. (1987), 'Econometric Estimation of Producers' Risk Attitudes', *American Journal of Agricultural Economics,* 69 (3), 509-22.

--- (2010), 'Asymmetry, Partial Moments, and Production Risk', *American Journal of Agricultural Economics,* 92 (5), 1294-309.

Asfaw, Solomon, et al. (2012a), 'Impact of modern agricultural technologies on smallholder welfare: Evidence from Tanzania and Ethiopia', *Food Policy,* 37 (3), 283-95.

Asfaw, Solomon, et al. (2012b), 'Poverty reduction effects of agricultural technology adoption: a micro-evidence from rural Tanzania', *Journal of Development Studies,* 48 (9), 1288-305.

Baland, Jean-Marie, Guirkinger, Catherine, and Mali, Charlotte (2011), 'Pretending to be poor: Borrowing to escape forced solidarity in Cameroon', *Economic Development and Cultural Change,* 60 (1), 1-16.

Bandiera, Oriana and Rasul, Imran (2006), 'Social networks and technology adoption in northern mozambique*', *The Economic Journal,* 116 (514), 869-902.

Bank, World (2008), *World development report* (Washington DC).

Barbier, Edward B (2010), 'Poverty, development, and environment', *Environment and Development Economics,* 15 (06), 635-60.

Barr, Abigail (2000), 'Social capital and technical information flows in the Ghanaian manufacturing sector', *Oxford Economic Papers,* 52 (3), 539-59.

Barrett, Christopher B and Arcese, Peter (1998), 'Wildlife harvest in integrated conservation and development projects: linking harvest to household demand, agricultural production, and environmental shocks in the Serengeti', *Land economics*, 449-65.

Barrett, Christopher B, Reardon, Thomas, and Webb, Patrick (2001), 'Nonfarm income diversification and household livelihood strategies in rural Africa: concepts, dynamics, and policy implications', *Food policy,* 26 (4), 315-31.

Barrett, Christopher B, Travis, Alexander J, and Dasgupta, Partha (2011), 'On biodiversity conservation and poverty traps', *Proceedings of the National Academy of Sciences,* 108 (34), 13907-12.

Battese, George Edward, Rambaldi, AN, and Wan, GH (1997), 'A stochastic frontier production function with flexible risk properties', *Journal of Productivity Analysis,* 8 (3), 269-80.

Becerril, Javier and Abdulai, Awudu (2010), 'The impact of improved maize varieties on poverty in Mexico: a propensity score-matching approach', *World Development,* 38 (7), 1024-35.

Bera, Anil K and Sharma, Subhash C (1999), 'Estimating production uncertainty in stochastic frontier production function models', *Journal of Productivity Analysis,* 12 (3), 187-210.

Berentsen, PBM, Kovacs, K, and Van Asseldonk, MAPM (2012), 'Comparing risk in conventional and organic dairy farming in the Netherlands: an empirical analysis', *Journal of dairy science,* 95 (7), 3803-11.

Berhane, Guush, et al. (2015), 'Formal and informal insurance: experimental evidence from Ethiopia', *2015 Conference, August 9-14, 2015, Milan, Italy* (International Association of Agricultural Economists).

Bernoulli, Daniel (1954), 'Exposition of a New Theory on the Measurement of Risk', *Econometrica,* 22 (1), 23-36.

Besley, Timothy and Case, Anne (1993), 'Modeling technology adoption in developing countries', *The American Economic Review*, 396-402.

Bezabih, Mintewab and Sarr, Mare (2012), 'Risk preferences and environmental uncertainty: Implications for crop diversification decisions in Ethiopia', *Environmental and Resource Economics,* 53 (4), 483-505.

Bezu, Sosina, et al. (2014), 'Impact of improved maize adoption on welfare of farm households in Malawi: a panel data analysis', *World Development,* 59, 120-31.

Bigsten, Arne and Shimeles, Abebe (2008), 'Poverty transition and persistence in Ethiopia: 1994–2004', *World Development,* 36 (9), 1559-84.

Binswanger, Hans P (1980), 'Attitudes toward risk: Experimental measurement in rural India', *American journal of agricultural economics,* 62 (3), 395-407.

Bramoullié, Yann and Kranton, Rachel (2007), 'Risk Sharing across Communities', *The American Economic Review,* 97 (2), 70-74.

Brick, Kerri and Visser, Martine (2015), 'Risk preferences, technology adoption and insurance uptake: A framed experiment', *Journal of Economic Behavior & Organization*.

Buschena, David E. (2003), 'Expected Utility Violations: Implications for Agricultural and Natural Resource Economics', *American Journal of Agricultural Economics,* 85 (5), 1242-48.

Chambers, Robert G. and Quiggin, John (1998), 'Cost Functions and Duality for Stochastic Technologies', *American Journal of Agricultural Economics,* 80 (2), 288-95.

Chavas, Jean-Paul and Holt, Matthew T. (1996), 'Economic Behavior Under Uncertainty: A Joint Analysis of Risk Preferences and Technology', *The Review of Economics and Statistics,* 78 (2), 329-35.

Chavas, Jean-Paul and Shi, Guanming (2015), 'An Economic Analysis of Risk, Management, and Agricultural Technology', *Journal of Agricultural and Resource Economics,* 40 (1), 63-79.

Chuang, Yating and Schechter, Laura (2015), 'Social Networks in Developing Countries', *Annual Review of Resource Economics,* 7.

Cilliers, Jacobus, Dube, Oeindrila, and Siddiqi, Bilal (2015), 'The white-man effect: How foreigner presence affects behavior in experiments', *Journal of Economic Behavior & Organization,* 118, 397-414.

Clarke, Daniel J. (2011), 'Insurance Design for Developing Countries, PhD Thesis', (Oxford).

Conley, Timothy G and Udry, Christopher R (2010), 'Learning about a new technology: Pineapple in Ghana', *The American Economic Review*, 35-69.

Conway, Gordon (1998), *The doubly green revolution: food for all in the twenty-first century* (Cornell University Press).

Dasgupta, Susmita, et al. (2005), 'Where is the poverty–environment nexus? Evidence from Cambodia, Lao PDR, and Vietnam', *World Development,* 33 (4), 617-38.
Dell, Melissa, Jones, Benjamin F., and Olken, Benjamin A. (2014), 'What Do We Learn from the Weather? The New Climate-Economy Literature', *Journal of Economic Literature,* 52 (3), 740-98.
Dercon, Stefan (1996), 'Risk, Crop Choice, and Savings: Evidence from Tanzania', *Economic Development and Cultural Change,* 44 (3), 485-513.
--- (2002), 'Income risk, coping strategies, and safety nets', *The World Bank Research Observer,* 17 (2), 141-66.
--- (2004), 'Growth and shocks: evidence from rural Ethiopia', *Journal of Development Economics,* 74 (2), 309-29.
--- (2008), 'Fate and fear: Risk and its consequences in Africa', *Journal of African Economies,* 17 (suppl 2), ii97-ii127.
Dercon, Stefan and Krishnan, Pramila (2000), 'Vulnerability, seasonality and poverty in Ethiopia', *The Journal of Development Studies,* 36 (6), 25-53.
Dercon, Stefan and Christiaensen, Luc (2011), 'Consumption risk, technology adoption and poverty traps: Evidence from Ethiopia', *Journal of development economics,* 96 (2), 159-73.
Dercon, Stefan and Porter, Catherine (2014), 'Live Aid Revisited: Long-Term Impacts of the 1984 Ethiopian Famine on Children', *Journal of the European Economic Association,* 12 (4), 927-48.
Dercon, Stefan, Hoddinott, John, and Woldehanna, Tassew (2005), 'Shocks and consumption in 15 Ethiopian villages, 1999-2004', *Journal of African Economies*.
Dercon, Stefan, et al. (2014), 'Offering rainfall insurance to informal insurance groups: Evidence from a field experiment in Ethiopia', *Journal of Development Economics,* 106, 132-43.
Di Falco, Salvatore and Perrings, Charles (2005), 'Crop biodiversity, risk management and the implications of agricultural assistance', *Ecological economics,* 55 (4), 459-66.
Di Falco, Salvatore and Chavas, Jean-Paul (2006), 'Crop genetic diversity, farm productivity and the management of environmental risk in rainfed agriculture', *European review of agricultural economics,* 33 (3), 289-314.
--- (2009), 'On crop biodiversity, risk exposure, and food security in the highlands of Ethiopia', *American Journal of Agricultural Economics,* 91 (3), 599-611.
Di Falco, Salvatore and Bulte, Erwin (2011), 'A dark side of social capital? Kinship, consumption, and savings', *Journal of Development Studies,* 47 (8), 1128-51.
Di Falco, Salvatore, Veronesi, Marcella, and Yesuf, Mahmud (2011), 'Does adaptation to climate change provide food security? A micro-perspective from Ethiopia', *American Journal of Agricultural Economics,* 93 (3), 829-46.
Duflo, Esther, Kremer, Michael, and Robinson, Jonathan (2008), 'How high are rates of return to fertilizer? Evidence from field experiments in Kenya', *The American economic review*, 482-88.
--- (2011), 'Nudging Farmers to Use Fertilizer: Theory and Experimental Evidence from Kenya', *American Economic Review,* 101 (6), 2350-90.
Eeckhoudt, Louis and Schlesinger, Harris (2005), 'Putting Risk in its Proper Place', *CESIFO Working Papers, Empirical and theoretical methods,* NO. 1462.
Ellsberg, Daniel (1961), 'Risk, ambiguity, and the Savage axioms', *The quarterly journal of economics*, 643-69.

Etner, Johanna, Jeleva, Meglena, and Tallon, Jean-Marc (2012), 'Decision theory under ambiguity', *Journal of Economic Surveys,* 26 (2), 234-70.

Fafchamps, Marcel (1992), 'Solidarity networks in preindustrial societies: Rational peasants with a moral economy', *Economic development and cultural change*, 147-74.

--- (2009), 'Vulnerability, Risk Management, and Agricultural Development', *CSAE Working Paper Series* (Centre for the Study of African Economies, University of Oxford).

Fafchamps, Marcel and Lund, Susan (2003), 'Risk-sharing networks in rural Philippines', *Journal of development Economics,* 71 (2), 261-87.

Fafchamps, Marcel and Gubert, Flore (2007), 'The formation of risk sharing networks', *Journal of Development Economics,* 83 (2), 326-50.

Fafchamps, Marcel, Udry, Christopher, and Czukas, Katherine (1998), 'Drought and saving in West Africa: are livestock a buffer stock?', *Journal of Development economics,* 55 (2), 273-305.

FAO, Food and Agriculture Organization of the United Nations (2009), 'How to Feed the World in 2050', (Rome).

Feder, Gershon, Just, Richard E, and Zilberman, David (1985), 'Adoption of agricultural innovations in developing countries: A survey', *Economic development and cultural change*, 255-98.

Finger, Robert (2014), 'Risk considerations in the economic assessment of low-input crop production techniques: an example from Swiss wheat production', *International Journal of Agricultural Resources, Governance and Ecology,* 10 (1), 63-77.

Fisher, Monica, et al. (2015), 'Drought tolerant maize for farmer adaptation to drought in sub-Saharan Africa: Determinants of adoption in eastern and southern Africa', *Climatic Change,* 133 (2), 283-99.

Foster, Andrew D and Rosenzweig, Mark R (1995), 'Learning by doing and learning from others: Human capital and technical change in agriculture', *Journal of political Economy*, 1176-209.

--- (2010), 'Microeconomics of technology adoption', *Annual Review of Economics,* 2.

Foudi, Sébastien and Erdlenbruch, Katrin (2011), 'The role of irrigation in farmers' risk management strategies in France', *European Review of Agricultural Economics*.

Gardebroek, Cornelis (2006), 'Comparing risk attitudes of organic and non-organic farmers with a Bayesian random coefficient model', *European Review of Agricultural Economics,* 33 (4), 485-510.

Gilboa, Itzhak (2009), *Theory of decision under uncertainty* (1: Cambridge university press Cambridge).

Gloede, Oliver, Menkhoff, Lukas, and Waibel, Hermann (2013), 'Shocks, individual risk attitude, and vulnerability to poverty among rural households in Thailand and Vietnam', *World Development*.

Groom, Ben, et al. (2008), 'The story of the moment: risk averse cypriot farmers respond to drought management', *Applied Economics,* 40 (3), 315-26.

Hardaker, J. B., et al. (2004), *Coping With Risk in Agriculture* (2 ed. edn.; Cambridge: CABI).

Harrison, Glenn W, Humphrey, Steven J, and Verschoor, Arjan (2010), 'Choice under uncertainty: evidence from Ethiopia, India and Uganda*', *The Economic Journal,* 120 (543), 80-104.

Hazell, Peter B. R. (1992), 'The appropriate role of agricultural insurance in developing countries', *Journal of International Development,* 4 (6), 567-81.

Hidalgo, F Daniel, et al. (2010), 'Economic determinants of land invasions', *The Review of Economics and Statistics,* 92 (3), 505-23.
Hill, Ruth Vargas (2009), 'Using stated preferences and beliefs to identify the impact of risk on poor households', *The Journal of Development Studies,* 45 (2), 151-71.
Hill, Ruth Vargas, Hoddinott, John, and Kumar, Neha (2013), 'Adoption of weather-index insurance: learning from willingness to pay among a panel of households in rural Ethiopia', *Agricultural Economics,* 44 (4-5), 385-98.
Hoddinott, John (2006), 'Shocks and their consequences across and within households in rural Zimbabwe', *The Journal of Development Studies,* 42 (2), 301-21.
Humphrey, Steven J and Verschoor, Arjan (2004), 'Decision-making under risk among small farmers in East Uganda', *Journal of African Economies,* 13 (1), 44-101.
Ibarra, Hector and Skees, Jerry (2007), 'Innovation in risk transfer for natural hazards impacting agriculture', *Environmental Hazards,* 7 (1), 62-69.
IPCC, Intergovernmental Panel on Climate Change (2007), *IPCC Fourth Assessment Report (AR4)* (Geneva).
Isham, Jonathan (2002), 'The effect of social capital on fertiliser adoption: Evidence from rural Tanzania', *Journal of African Economies,* 11 (1), 39-60.
Jaenicke, Edward C, Frechette, Darren L, and Larson, James A (2003), 'Estimating production risk and inefficiency simultaneously: an application to cotton cropping systems', *Journal of Agricultural and Resource Economics*, 540-57.
Jakiela, Pamela and Ozier, Owen (2016), 'Does Africa Need a Rotten Kin Theorem? Experimental Evidence from Village Economies', *The Review of Economic Studies,* 83 (1), 231-68.
Just, David R., Khantachavana, Sivalai V., and Just, Richard E. (2010), 'Empirical Challenges for Risk Preferences and Production', *Annual Review of Resource Economics,* 2 (1), 13-31.
Just, Richard E and Zilberman, David (1983), 'Stochastic structure, farm size and technology adoption in developing agriculture', *Oxford Economic Papers*, 307-28.
Just, Richard E. and Pope, Rulon D. (1978), 'Stochastic specification of production functions and economic implications', *Journal of Econometrics,* 7 (1), 67-86.
Kahneman, Daniel and Tversky, Amos (1979), 'Prospect Theory: An Analysis of Decision under Risk', *Econometrica,* 47 (2), 263-91.
Karlan, Dean, et al. (2014), 'Agricultural Decisions after Relaxing Credit and Risk Constraints', *The Quarterly Journal of Economics,* 129 (2), 597-652.
Kassie, Menale, Jaleta, Moti, and Mattei, Alessandra (2014), 'Evaluating the impact of improved maize varieties on food security in Rural Tanzania: Evidence from a continuous treatment approach', *Food Security,* 6 (2), 217-30.
Kazianga, Harounan and Udry, Christopher (2006), 'Consumption smoothing? Livestock, insurance and drought in rural Burkina Faso', *Journal of Development Economics,* 79 (2), 413-46.
Khonje, Makaiko, et al. (2015), 'Analysis of Adoption and Impacts of Improved Maize Varieties in Eastern Zambia', *World Development,* 66, 695-706.
Kim, Kwansoo, et al. (2014), 'Rice, irrigation and downside risk: a quantile analysis of risk exposure and mitigation on Korean farms', *European Review of Agricultural Economics,* 41 (5), 775-815.
Kinsey, Bill, Burger, Kees, and Gunning, Jan Willem (1998), 'Coping with drought in Zimbabwe: Survey evidence on responses of rural households to risk', *World Development,* 26 (1), 89-110.

Knight, Frank Hyneman (1921), *Risk, Uncertainty and Profit* (Hart, Schaffner & Marx prize essays; Boston: Houghton Mifflin).
Knight, John, Weir, Sharada, and Woldehanna, Tassew (2003), 'The role of education in facilitating risk-taking and innovation in agriculture', *The Journal of Development Studies,* 39 (6), 1-22.
Koundouri, Phoebe, et al. (2009), 'The effects of EU agricultural policy changes on farmers' risk attitudes', *Eur Rev Agric Econ*, jbp003.
Kumbhakar, Sabul C (1993), 'Production risk, technical efficiency, and panel data', *Economics Letters,* 41 (1), 11-16.
Kumbhakar, Subal and Tsionas, Efthymios (2010), 'Estimation of production risk and risk preference function: a nonparametric approach', *Annals of Operations Research,* 176 (1), 369-78.
Kumbhakar, Subal C. (2002), 'Specification and Estimation of Production Risk, Risk Preferences and Technical Efficiency', *American Journal of Agricultural Economics,* 84 (1), 8-22.
Kumbhakar, Subal C. and Tveterås, Ragnar (2003), 'Risk Preferences, Production Risk and Firm Heterogeneity*', *The Scandinavian Journal of Economics,* 105 (2), 275-93.
Lambrecht, Isabel, et al. (2014), 'Understanding the process of agricultural technology adoption: mineral fertilizer in eastern DR Congo', *World Development,* 59, 132-46.
Lence, Sergio H. (2009), 'Joint Estimation of Risk Preferences and Technology: Flexible Utility or Futility?', *American Journal of Agricultural Economics,* 91 (3), 581-98.
Little, Peter D, et al. (2001), 'Avoiding disaster: diversification and risk management among East African herders', *Development and Change,* 32 (3), 401-33.
Lobell, David B, Schlenker, Wolfram, and Costa-Roberts, Justin (2011a), 'Climate trends and global crop production since 1980', *Science,* 333 (6042), 616-20.
Lobell, David B, et al. (2011b), 'Nonlinear heat effects on African maize as evidenced by historical yield trials', *Nature Climate Change,* 1 (1), 42-45.
Love, H. Alan and Buccola, Steven T. (1991), 'Joint Risk Preference-Technology Estimation with a Primal System', *American Journal of Agricultural Economics,* 73 (3), 765-74.
Machina, Mark J. (1987), 'Choice Under Uncertainty: Problems Solved and Unsolved', *The Journal of Economic Perspectives,* 1 (1), 121-54.
Maertens, Annemie and Barrett, Christopher B (2013), 'Measuring social networks' effects on agricultural technology adoption', *American Journal of Agricultural Economics,* 95 (2), 353-59.
Mao, James C. T. (1970), 'Survey of Capital Budgeting: Theory and Practice', *The Journal of Finance,* 25 (2), 349-60.
Mathenge, Mary K, Smale, Melinda, and Olwande, John (2014), 'The impacts of hybrid maize seed on the welfare of farming households in Kenya', *Food Policy,* 44, 262-71.
Mendelsohn, Robert, Nordhaus, William D, and Shaw, Daigee (1994), 'The impact of global warming on agriculture: a Ricardian analysis', *The American economic review*, 753-71.
Mendola, Mariapia (2007), 'Agricultural technology adoption and poverty reduction: A propensity-score matching analysis for rural Bangladesh', *Food policy,* 32 (3), 372-93.
Menezes, C., Geiss, C., and Tressler, J. (1980), 'Increasing Downside Risk', *The American Economic Review,* 70 (5), 921-32.
Minten, Bart and Barrett, Christopher B (2008), 'Agricultural technology, productivity, and poverty in Madagascar', *World Development,* 36 (5), 797-822.
Morduch, Jonathan (1999), 'Between the state and the market: Can informal insurance patch the safety net?', *The World Bank Research Observer,* 14 (2), 187-207.

Morris, J, Mills, J, and Crawford, IM (2000), 'Promoting farmer uptake of agri-environment schemes: the Countryside Stewardship Arable Options Scheme', *Land Use Policy,* 17 (3), 241-54.

Moser, Christine M and Barrett, Christopher B (2006), 'The complex dynamics of smallholder technology adoption: the case of SRI in Madagascar', *Agricultural Economics,* 35 (3), 373-88.

Mosley, Paul and Verschoor, Arjan (2005), 'Risk attitudes and the 'vicious circle of poverty'', *The European journal of development research,* 17 (1), 59-88.

Munshi, Kaivan (2004), 'Social learning in a heterogeneous population: technology diffusion in the Indian Green Revolution', *Journal of Development Economics,* 73 (1), 185-213.

Nastis, Stefanos A, Michailidis, Anastasios, and Mattas, Konstadinos (2013), 'Crop biodiversity repercussions of subsidized organic farming', *Land use policy,* 32, 23-26.

Peerlings, Jack and Polman, Nico (2009), 'Farm choice between agri-environmental contracts in the European Union', *Journal of Environmental Planning and Management,* 52 (5), 593-612.

Pope, Rulon D. and Just, Richard E. (1991), 'On Testing the Structure of Risk Preferences in Agricultural Supply Analysis', *American Journal of Agricultural Economics,* 73 (3), 743-48.

Porter, Catherine (2012), 'Shocks, consumption and income diversification in rural Ethiopia', *The Journal of Development Studies,* 48 (9), 1209-22.

Pratt, John W. (1964), 'Risk Aversion in the Small and in the Large', *Econometrica,* 32 (1/2), 122-36.

Reardon, Thomas, et al. (2000), 'Effects of non-farm employment on rural income inequality in developing countries: an investment perspective', *Journal of agricultural economics,* 51 (2), 266-88.

Rosenzweig, Mark R and Binswanger, Hans P (1993), 'Wealth, weather risk and the composition and profitability of agricultural investments', *The economic journal*, 56-78.

Rothschild, Michael and Stiglitz, Joseph E. (1970), 'Increasing Risk: I. A Definition', *Journal of Economic Theory,* 2, 225-43.

--- (1971), 'Increasing Risk II: Its economic consequences', *Journal of Economic Theory,* 3 (1), 66-84.

Saastamoinen, Antti (2013), 'Heteroscedasticity or Production Risk? A Synthetic View', *Journal of Economic Surveys*.

Saha, Atanu, Shumway, C. Richard, and Talpaz, Hovav (1994), 'Joint Estimation of Risk Preference Structure and Technology Using Expo-Power Utility', *American Journal of Agricultural Economics,* 76 (2), 173-84.

Sandmo, Agnar (1971), 'On the Theory of the Competitive Firm Under Price Uncertainty', *The American Economic Review,* 61 (1), 65-73.

Santos, Paulo and Barrett, Christopher B (2011), 'Persistent poverty and informal credit', *Journal of Development Economics,* 96 (2), 337-47.

Savage, L. J. (1954), *The Foundations of Statistics* (New York: John Wiley).

Schlenker, Wolfram and Roberts, Michael J (2009), 'Nonlinear temperature effects indicate severe damages to US crop yields under climate change', *Proceedings of the National Academy of sciences,* 106 (37), 15594-98.

Schlenker, Wolfram and Lobell, David B (2010), 'Robust negative impacts of climate change on African agriculture', *Environmental Research Letters,* 5 (1), 014010.

Serra, Teresa, Zilberman, David, and Gil, José M (2008), 'Differential uncertainties and risk attitudes between conventional and organic producers: the case of Spanish arable crop farmers', *Agricultural Economics,* 39 (2), 219-29.
Shiferaw, Bekele, et al. (2014), 'Adoption of improved wheat varieties and impacts on household food security in Ethiopia', *Food Policy,* 44, 272-84.
Simtowe, Franklin, et al. (2006), 'Can Risk-aversion Toward Fertilizer Explain Part of the Non-adoption Puzzle for Hybrid Maize? Empirical Evidence from Malawi', *Journal of Applied Sciences,* 6 (7), 1490-98.
Smale, Melinda and Mason, Nicole (2014), 'Hybrid seed and the economic well-being of smallholder maize farmers in Zambia', *Journal of Development Studies,* 50 (5), 680-95.
Smale, Melinda, et al. (1998), 'The Contribution of Genetic Resources and Diversity to Wheat Production in the Punjab of Pakistan', *American Journal of Agricultural Economics,* 80 (3), 482-93.
Starmer, Chris (2000), 'Developments in non-expected utility theory: The hunt for a descriptive theory of choice under risk', *Journal of economic literature*, 332-82.
Suri, Tavneet (2011), 'Selection and comparative advantage in technology adoption', *Econometrica,* 79 (1), 159-209.
Tack, Jesse, Harri, Ardian, and Coble, Keith (2012), 'More than mean effects: Modeling the effect of climate on the higher order moments of crop yields', *American Journal of Agricultural Economics*, aas071.
Tack, Jesse B and Ubilava, David (2013), 'The effect of El Niño Southern Oscillation on US corn production and downside risk', *Climatic change,* 121 (4), 689-700.
Tanaka, Yuki and Munro, Alistair (2014), 'Regional variation in risk and time preferences: Evidence from a large-scale field experiment in rural Uganda', *Journal of African Economies,* 23 (1), 151-87.
Townsend, Robert M. (1994), 'Risk and Insurance in Village India', *Econometrica,* 62 (3), 539-91.
Tversky, Amos and Kahneman, Daniel (1974), 'Judgment under Uncertainty: Heuristics and Biases', *Science,* 185 (4157), 1124-31.
United Nations, Department of Economic and Social Affairs, Population Division (2015), 'World Population Prospects: The 2015 Revision, Key Findings and Advance Tables.', *Working Paper,* ESA/P/WP.241.
Van den Broeck, Katleen and Dercon, Stefan (2011), 'Information flows and social externalities in a Tanzanian banana growing village', *The journal of development studies,* 47 (2), 231-52.
von Neumann, John and Morgenstern, Oscar (1944), *Theory Games and Economic Behavior* (Princeton: Princeton University Press).
Wakker, Peter P (2010), *Prospect theory: For risk and ambiguity* (Cambridge University Press).
Wang, Hung-Jen and Schmidt, Peter (2002), 'One-step and two-step estimation of the effects of exogenous variables on technical efficiency levels', *journal of Productivity Analysis,* 18 (2), 129-44.
Weir, Sharada and Knight, John (2000), *Adoption and diffusion of agricultural innovations in Ethiopia: the role of education* (University of Oxford, Institute of Economics and Statistics, Centre for the Study of African Economies).
Wekesa, E, et al. (2003), *Adoption of maize production technologies in the coastal lowlands of Kenya* (CIMMYT).

Welch, Jarrod R, et al. (2010), 'Rice yields in tropical/subtropical Asia exhibit large but opposing sensitivities to minimum and maximum temperatures', *Proceedings of the National Academy of Sciences,* 107 (33), 14562-67.

Wynn, Gerard, Crabtree, Bob, and Potts, Jacqueline (2001), 'Modelling farmer entry into the environmentally sensitive area schemes in Scotland', *Journal of agricultural economics,* 52 (1), 65-82.

Yang, Dean and Choi, HwaJung (2007), 'Are remittances insurance? Evidence from rainfall shocks in the Philippines', *The World Bank Economic Review,* 21 (2), 219-48.

Yesuf, Mahmud and Bluffstone, Randall A (2009), 'Poverty, risk aversion, and path dependence in low-income countries: Experimental evidence from Ethiopia', *American Journal of Agricultural Economics,* 91 (4), 1022-37.

Yilma, Zelalem, et al. (2014), 'Coping with shocks in rural Ethiopia', *Journal of Development Studies,* 50 (7), 1009-24.

Zeller, Manfred, Diagne, Aliou, and Mataya, Charles (1998), 'Market access by smallholder farmers in Malawi: Implications for technology adoption, agricultural productivity and crop income', *Agricultural Economics,* 19 (1), 219-29.

Zimmerman, Frederick J and Carter, Michael R (2003), 'Asset smoothing, consumption smoothing and the reproduction of inequality under risk and subsistence constraints', *Journal of Development Economics,* 71 (2), 233-60.

2 A Simple Framework for the Estimation of Climate Risk Exposure

Xavier Vollenweider[6]

Abstract

This article introduces a new framework to estimate climate risk exposure at the household level with the standardized precipitation evapotranspiration index (SPEI) as its building block. The great advantage of using the SPEI is in knowing that it is distributed as a standard normal distribution. We have hence a proxy variable for the climate with a known distribution. Once the conditional expectation of consumption has been estimated as a function of SPEI, the probability density function of expected consumption conditional on SPEI can be derived by a change of variables. We use this probability density function as our measure of climate risk exposure. Furthermore, the approach is simple enough to accommodate quantile regressions and hence offers the opportunity to broaden the scope of the analysis to different categories of the population. Lastly, it offers a direct estimate of the central moments of the climate risk exposure function via the regression estimates on the SPEI (variance, skewness, kurtosis). It circumvents hence the use of the residuals as done in traditional model of production risk analysis. The methodology is illustrated with a case study on Ethiopia, combining data from the Ethiopian Rural Household Survey (ERHS) with climate data (African Rainfall Climatology Version 2 dataset and Climate Prediction Centre Global Land Surface Air Temperature Analysis, GHCN+CAMS, NOAA 2001). The results show notably that households located at low altitudes are the most exposed to climate risk.

[6] Department of Geography and Environment, London School of Economics and Political Science, United Kingdom. Contact author: Xavier Vollenweider: X.Y.Vollenweider@lse.ac.uk. I gratefully acknowledge the support from the Walsh fellowship and the Rural Economy and Development Programme of Teagasc, the Agriculture and Food Development Authority of the Republic of Ireland. I would like also to extend my thanks to Elvezio Ronchetti, Ben Groom, Kelvin Balcombe and Steve Gibbons for comments on earlier drafts of this paper.

2.1 Introduction

The seminal paper of Sandmo (1971) showing that risk leads to underinvestment and underproduction contributed to establishing the economics of production under uncertainty as an important research stream in economics, with agriculture as one of its favourite case studies. If production risk is a major topic in the agricultural economics literature, it is probably because 'the most singular aspect of agricultural production is its randomness' (Chambers and Quiggin 1998). The main framework for production risk estimation is based on the stochastic production analysis of Just and Pope (1978) and Antle (1983). These models, and their later extensions to skewness and efficiency analysis (Di Falco and Chavas 2006; Kumbhakar and Tveterås 2003), have been the backbone of a large number of studies. They have been applied to the estimation of risk preferences, and efficiency (e.g. John M Antle 1987; Koundouri et al. 2009; Love and Buccola 1991), to estimate the role of biodiversity as a risk mitigating option (e.g. Di Falco and Chavas 2006; Di Falco and Chavas 2009; Smale et al. 1998) and to water resource management (e.g. Groom et al. 2008). See Saastamoinen (2013) for an recent and synthetic literature review.

Although the existing estimation framework is appropriate for estimating *short-term* production risk, the estimation of climate exposure is more elusive: climate risk in the classical framework is lumped into the larger category of production risk; a catch-all term covering plant and animal diseases, pests, mushrooms, damage caused by animals as well as droughts and floods. Two main reasons can explain this gap in the literature. First, when the foundations of the stochastic production analysis framework were laid, i.e. the beginning of the 1980s, climate change was not yet on the political agenda. Second, weather data were not widely available in the 1980s and geographical information system (GIS) software was still the realm of a few specialists.

Following the Rio Declaration on Environment and Development (1992), the emergence of climate change and climate adaptation as a serious challenge to policymakers at both national and international levels has highlighted the need for a precise estimation of household climate risk exposure. Furthermore, anyone can nowadays access daily satellite and weather station precipitation as well as temperature data over several decades, and link them with microeconomic data thanks to GIS software (e.g. Quantum GIS[7], R[8]). Hence, a new methodology utilizing this climate data bonanza and answering these policy needs is required.

[7] Quantum GIS Development Team (2013). Quantum GIS Geographic Information System. Open Source Geospatial Foundation Project. http://qgis.osgeo.org.

[8] R Core Team (2013). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.

So far, the focus has been on estimating the production risk of the average household. Indeed, the main tool to investigate changes in other parts of a population distribution, i.e. quantile regression analysis (Koenker and Bassett 1978), was still a novelty at the time of the pioneering work of Just and Pope (1978). It is, however, of interest to know how climate exposure varies between poor and rich households or if a particular development policy is effective at decreasing climate exposure among poorer parts of a population. Standard quantile regression routines are now widely available on common statistical software (e.g. STATA) and their extensions to panel data, still an active field of research, are readily available via the R CRAN project, for instance. The new methodology should hence be simple enough to accommodate quantile regressions in order to distinguish climate exposure in different categories of the population.

The methodology proposed in the present article is built on the use of standardized measures of weather. The Standardized Precipitation Index (SPI), first introduced by McKee et al. (McKee et al. 1993, 1995), is a locally and frequency based characterization of precipitation levels. Guttman (1998, 1999) widely contributed to its popularisation by showing some of its key advantages over the Palmer Drought Severity Index (Palmer 1965), the index of choice at the time.[9] The SPI allows the comparison of hydrological conditions across space and time (Hayes et al. 1999), is flexible enough to consider different kind of droughts (e.g. hydrological conditions on monthly scales affecting agriculture, or at yearly scales affecting large-scale water management), is simple and tractable, and is parsimonious in terms of data requirements.

Note that climate change affects both changes in precipitation and temperature. Vicente-Serrano et al. (2010) have proposed the SPEI in order to take into account the influence of temperature on hydrological conditions. Its statistical conception and properties are essentially the same as the SPI. However, in the case of the SPEI, it is the difference between precipitation and potential evapotranspiration, i.e. the net balance of water, which is standardized. As both temperature and precipitation have an impact on agricultural production and the livelihood of rural populations, and since the SPEI is more sensible in the context of climate change, it was preferred as the standardized measure of weather.

The use of the SPEI offers the opportunity to easily characterize average production or consumption under locally and frequency-defined weather scenarios. As the framework is very simple, it can easily

[9] The Palmer Drought Severity Index (PDSI) is based on a water balance equation taking into account precipitation, moisture supply, runoff and evaporation demand at the surface level. According to Vicente-Serrano et al. (2010), although some of the weaknesses of the PSDI have been solved by Wells et al. (2004), the main weakness of the PDSI identified by Guttman (1998) has not been addressed: the fixed temporal scale between 9 to 12 months and the fact that PDSI values are affected by conditions up to four years in the past.

be extended to quantile regressions in order to broaden the scope of analysis to households at different quantiles of the population distribution. In order to control for unobserved heterogeneity, we rely on penalized quantile fixed effects quantile regressions as proposed by Koenker (2004).

Once the climate risk exposure has been estimated, a vulnerability index is needed to summarize the information. We rely on three indices: (1) poverty risk, (2) expected shortfall, and (3) relative risk premium. We apply the proposed methodology to the consumption level of rural households in Ethiopia with data from the Ethiopian Rural Household Survey (ERHS)[10], a panel dataset with seven rounds conducted between 1989 and 2009, including more than 1,200 households. The climate data come from the African Rainfall Climatology Version 2 dataset and the Climate Prediction Center Global Land Surface Air Temperature Analysis dataset (GHCN+CAMS, NOAA 2001). All datasets used in the present study are freely available online.

Section 2.2 presents the estimation framework starting with a brief review of the classical estimation framework of production risk analysis (2.2.1), following with the presentation of the SPEI (2.2.2) and the derivation of the climate consumption model (2.2.3). The vulnerability indices and the estimation strategy are then presented (2.2.4 and 2.2.5 respectively). Section 2.3 presents the data and results are discussed in section 2.4. Section 2.5 concludes.

2.2 Estimation framework

2.2.1 Production risk analysis

The classical risk estimation methodology was developed when climate and weather data were not widely available. The emphasis was hence on *production risk*, a catch term for drought, flood, pest and animal diseases. In other words, production risk was viewed as all factors affecting production which are not under the farmer's control, oscillating randomly from year to year and not related to market risk (e.g. inputs and outputs price volatility); resources risk (e.g. fertilizers, seeds and labour supply shocks), institutional risk (e.g. changes in policy), financial risk (e.g. changes in the interest rates

[10] The ERHS data have been made available by the Economics Department, Addis Ababa University (Economics /AAU); the Centre for the Study of African Economies, University of Oxford (CSAE); and the International Food Policy Research Institute (IFPRI). Funding for data collection was provided by the Economic and Social Research Council (ESRC), the Swedish International Development Agency (SIDA) and the United States Agency for International Development (USAID); the preparation of the public release version of these data was supported, in part, by the World Bank. AAU, CSAE, IFPRI, ESRC, SIDA, USAID and the World Bank are not responsible for any errors in these data or for their use or interpretation.

charged on the debt of the farm), personal risk (e.g. health issues, accidents), and asset risks (thefts or fire damages to buildings, machinery and livestock) (Hardaker et al. 2004; Hazell 1992). Note that financial risk, personal risk, and asset risk are rarely controlled for in applied studies and hence are lumped into production risk.

Furthermore, the framework was designed to disentangle the impact of different inputs on production risk exposure. The impact of weather risk on the production process was hence not the main concern. Most studies in the literature on poverty traps have addressed the question of weather shocks and weather risk impact on consumption either by including a dummy variable equal to one if the household was exposed to extreme events or by using another weather risk index. In the latter case, the most popular weather risk measure has been rainfall variability, captured by the variance or the intra-year coefficient of variation. However, such measures are likely to introduce unobserved heterogeneity bias if the sample overlaps different weather regimes. For instance, a great level of intra-year variation might be a characteristic of a particular weather regime and hence should not count as risk, while in another weather regime such variation would indeed imply erratic rainfalls. Dercon and Christansen (2011) use lower quantiles of the sample's rain distribution to characterize weather shocks. This approach is the closest to the one introduced in the present paper.

The goal of risk estimation could be summarized as the estimation of the different central moments of the probability distribution of production. The first central moment is the mean, i.e. the expected output or yield. The second moment, i.e. the variance, is a measure of the dispersion of the possible production levels. For instance, a farmer expecting a yield between 200 kg/ha and 4,000 kg/ha would have a higher variance than a farmer expecting a yield between 1,800 kg/ha and 2,200 kg/ha. Variance has hence been one of the first measures of risk. The third moment, summarized by the skewness, is a measure of the asymmetry of possible yields. Negative skewness implies that expected yield is lower than the most likely one and that if bad and good harvests with the same probability are compared, the bad harvest will cost more than the good one could have yield. It is hence often interpreted as a measure of downside risk.

The key insight of Just and Pope (1978) was to split the production function into a deterministic part and a stochastic part, allowing inputs to be risk-increasing, risk-neutral or risk-decreasing. The production, y, is specified as follows:

$$y = f(\boldsymbol{x}, \boldsymbol{\beta}) + h(\boldsymbol{x}, \boldsymbol{\gamma})^{1/2}\varepsilon \tag{1}$$

where $f(\boldsymbol{x}, \boldsymbol{\beta})$ is the deterministic production function, $\boldsymbol{x}$ a set of inputs, $\boldsymbol{\beta}$ a set of parameters to be estimated, $h(\boldsymbol{x}, \boldsymbol{\gamma})$ the risk function, with parameters $\boldsymbol{\gamma}$ to be estimated, and ε a random noise identically and independently distributed (*iid*) according to a standard normal distribution.

Antle (1983) showed that the Just and Pope approach restricts the effect of inputs across variance and higher moments. He proposed the so-called 'flexible moment-based approach' where the central probability moments (i.e. mean, variance etc.) are directly specified:

$$\mu_1(x, \beta_1) = \int y f(y|x) dy \tag{2}$$

$$\mu_i(x, \beta_i) = \int (y - \mu_1)^i f(y|x) dy \qquad i \geq 2 \tag{3}$$

where β_i relates the input x to the moment μ_i. This approach relaxes any cross-moments restrictions: the inputs' elasticity with respect to variance does not restrict their elasticity with respect to higher moments. The different moments can be estimated using a feasible generalized least square estimator (FGLS). The first step is hence to estimate a classical production function with FGLS, the residuals of which are then put to the square and to the cube to estimate the variance and skewness function. The predicted values of this set of three regressions are respectively the mean, variance and skewness of the conditional distribution of each farmer's production.

A limitation of these approaches is that they are highly parametric. Indeed, specification errors in the first moment, respectively Equations (1) and (2), cascade across the whole model, directly affecting the estimation of the higher moments.

A popular solution is to choose a flexible functional form such as the translog function, which corresponds to a second order Taylor approximation around the mean of the true production function (e.g. Greene 2003). Although mathematically appealing, the translog functions are notoriously hard to estimate with a sample of a few hundred observations (the usual sample size of rural household surveys): the set of covariates enters the function multiple times — in level, square and through the series of interaction terms — giving rise to important multicollinearity issues.[11] It is hence difficult to obtain statistically significant estimates and no test provides an objective criterion to select which covariates to retain. Full information maximum likelihood estimation and general method of moments provide more efficient results, although issues persist. As Kumbhakar and Tveteras (2003) note: '[the] idea of dropping insignificant variables is not pursued […] due to several problems. First, it destroys

[11] For instance, a production function with four explanatory variables, say labour, fertilizer, land and capital, implies fourteen parameters to estimate.

the flexibility of the mean output function. Second, dropping one insignificant variable caused other insignificant (significant) variables to be significant (insignificant) due to high multicollinearity (which is always present in flexible functions) and the use of a system approach. Furthermore, we found no natural order to select variables for exclusion in the present model'. Therefore, although the conditional expectation might well fit on average, marginal effects are difficult to ascertain.

Recently, quantile regressions started to attract interest in the microeconomics of risk literature (as across most applied statistical disciplines). The first author to mention the possible application of quantile regressions to production risk analysis is probably Charles B. Moss (2010), and the first to propose an estimation framework of production risk based on quantile regressions were Kim et al. (2014).

2.2.2 Definition of the climate variable

Meteorologists have struggled to give a definition of drought general enough to be comparable across areas and time: light rains in the middle of the rainy season might be the first sign of an incoming drought in a given area, while the same level of precipitation can be considered as totally normal at other times of the year or in another area. The standardized precipitation index (SPI) addresses precisely these kinds of issues. The SPI is a localized and statistical measure of precipitation. It offers a comparable index across times and regions. Indeed, it is based indeed on local frequency: given a series of cumulative local monthly precipitation over an extended period (30 years is deemed acceptable), probability functions are fitted on each monthly distribution. Most commonly, a gamma distribution is fitted with a maximum likelihood estimator and then standardized.

The SPI is symmetrically distributed around zero, a value of zero representing normal conditions, whilst below and above zero values represent dry and wet conditions respectively, with values between -0.5 and 0.5 considered as nearly normal. Although the SPI is theoretically unbounded, values below -3 and above 3 are extremely rare as they occur with a probability of 0.1 %. Assuming that weather events are identically and independently distributed, catastrophic droughts and floods can be defined as SPI values above and below $\pm$ 2.3, i.e. a drought or flood with a return period of 100 years (Guttmann 1999). Values above and below $\pm$ 1.9 can also be considered as extreme events as they have a return period of 35 years.

Recently, Vicente-Serrano et al. (2010) have proposed focusing on the net balance of water in order to take climate change into account. The intuition is the same, the only change being that it is not precipitation, but the difference between precipitation and evapotranspiration that is standardized.

The calculation of the SPEI has four main steps. We follow here the presentation of Vicente-Serrano et al. (2010). The first step consists in computing potential evapotranspiration (PET), i.e. the *demand* of water in the hydrological process. The simplest PET index, used in the present study, is the Thornthwaite Index (1948): it requires only the temperature and the latitude at which the data have been gathered. The derivation of this can be found in appendix 2.6.

The second step consists in computing the 'water balance' for a given month, say July, at time t, i.e. the difference between precipitation and evapotranspiration:

$$D_t = P_t - PET_t \tag{4}$$

where D_t, P_t and PET_t are respectively the water balance, the precipitation and the potential evapotranspiration measured in millimetres. A positive value for D_t implies at time t a water surplus and a negative one implies a water deficit. The step is then repeated on each month of July for which data are available in order to obtain a time series of the net balance of water in July over the last 30 years for instance.

The third step consists of fitting a distribution $F(D)$ on times series of D_t gathered over the sample period. The longer the period, the better is the distribution fit, but 30 years of data, i.e. 30 observations of D_t, is deemed acceptable. Several candidate distributions were investigated by Beguería and Vicente-Serrano (2013): Pearson III, Lognormal, Log-logistic and General Extreme Value. As all the investigated distributions fit well with empirical probabilities, a selection is made based on their behaviour at extreme values. Following the latter criterion, the log-logistic distribution is preferred and its parameters are estimated with the unbiased probability weighted moments method (Beguería and Vicente-Serrano 2013).

The last step consists of obtaining the SPEI values which are defined as the inverse of $F(D)$ once $F(D)$ is standardized. Beguería and Vicente-Serrano (2013) use the formula 26.2.23 in Abramowitz and Stegun (1972):

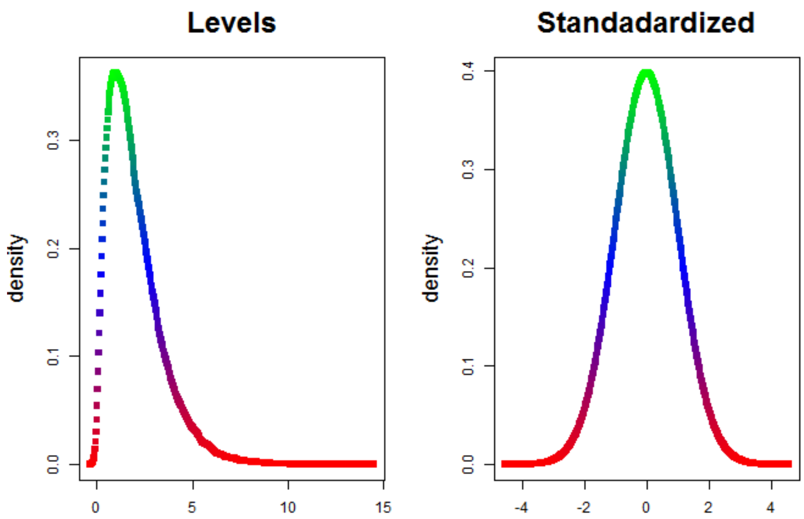
$$SPEI_p = W - \frac{C_0 + C_1 W + C_2 W^2}{1 + d_1 W + d_2 W^2 + d_3 W^3} \tag{5}$$

where p is the probability that D_t exceeds a given value $\overline{D}_t$ and is given by $F(D)$, $C_0 = 2.515517$, $C_1 = 0.802853$, $C_2 = 0.010328$, $d_1 = 1.432788$, $d_2 = 0.189269$, $d_3 = 0.001308$ and W is given by:

$$W = \sqrt{\ln\left(\frac{1}{p^2}\right)} = \sqrt{-2\ln p} \tag{6}$$

The approximation is valid for $p \leq 0.5$. For $p > 0.5$, p is replaced by $1 - p$ and the sign of the resulting $SPEI_p$ is inversed.

Figure 2.1: Standardization of the net balance of water



The process is illustrated with simulated data in Figure 2.1. On the left-hand side, we plot the density of net balance of water simulated according to a gamma distribution. The simulated data is then standardised with the formula in equation (5) in order to obtain the normal density plotted on the right-hand side, i.e. the SPEI index. The colour shading indicates the frequency of net balance of water's values, from very rare (dark red) to very frequent (bright green).

2.2.3 Climate consumption model

The great advantage of using the SPEI is in knowing that it is distributed as a standard normal distribution. We have hence a proxy variable for the climate with a known distribution[12]. Once the conditional expectation of consumption has been estimated as a function of SPEI, the probability density function of expected consumption conditional on SPEI can be derived by a change of variables. We use this probability density function as our measure of climate risk exposure. We present below how we recover climate risk exposure based on the relationship between consumption and SPEI.

[12] Following the terminology used in new climate-economy literature, we define the word 'climate' as the distribution of all possible weather events (Dell et al. 2014). We reserve the word weather of a realisation of the climate, i.e. a random draw from the distribution of all possible weather events.

Let us assume that the relation between the SPEI and consumption is defined as follows:

$$c = g(S) \tag{7}$$

where c is consumption, S is the SPEI with probability density function $f_s(S)$, i.e. a standard normal density function, and $g(S)$ is a monotonic and increasing function.

We can compute the probability density function of c by a change of variables as (e.g. Casella and Berger 2002):

$$f_c(c) = f_s(g^{-1}(c)) \left| \frac{dg^{-1}(c)}{dc} \right| \tag{8}$$

where $g^{-1}(c)$ is the inverse function of $g(S)$ and f_s is the probability density function of SPEI, i.e. a standard normal density function as shown in equation (5) and (6).

In the case of a non-monotonous function, we have (e.g. Casella and Berger 2002):

$$f_c(c) = \sum_{k=1}^{n(c)} \left| \frac{dg_k^{-1}(c)}{dc} \right| f_S\left(g_k^{-1}(c)\right) \tag{9}$$

where $g_k^{-1}(c) = S$ is the inverse function of $g(\,)$, $n(c)$ is the number of k solutions to $g_k^{-1}(c)$.

For the sake of example, let us assume that consumption is a linear function of SPEI:

$$c = g(S) = 2S \tag{10}$$

so that the $g^{-1}(c) = c/2$ and $dg^{-1}\,(c)/dc = 1/2$. Hence, the probability function of y can be expressed as:

$$f_c(c) = 0.5 f_s(\frac{c}{2}) \tag{11}$$

We illustrate the change of variables in Figure 2.2 with a simulation. We start by generating a sample of 1000 observations for S equally spaced over [-3, 3]. We then compute c according to (10) as shown by thick black line in plot a in Figure 2.2. We also plot f_s on its left axis (grey dashed line, plot a of Figure 2.2). Applying the formula in equation (8), we obtain the density function f_c and plot it in graph 1.b. Intuitively, each consumption level on the thick black line of c is weighted according to the likelihood of the corresponding S value shown in grey.

Many of the studies in the 'new climate-economy literature' use a non-linear specification of climate with a preference for the quadratic specification (e.g. Hidalgo et al. 2010; Lobell et al. 2011a; Lobell et al. 2011b; Schlenker and Roberts 2009; Schlenker and Lobell 2010). For instance, temperature increases yield up to a point where an increase in temperature has an adverse impact on yield, each additional degree decreasing it. We can expect the same to be true for consumption in rural areas of Sub-Saharan Africa, where weather conditions have a large effect on consumption (see section 1.4 of the literature review for a survey of the link between weather shocks and poverty).

We therefore repeat the example above, specifying this time consumption as a quadratic function of SPEI:

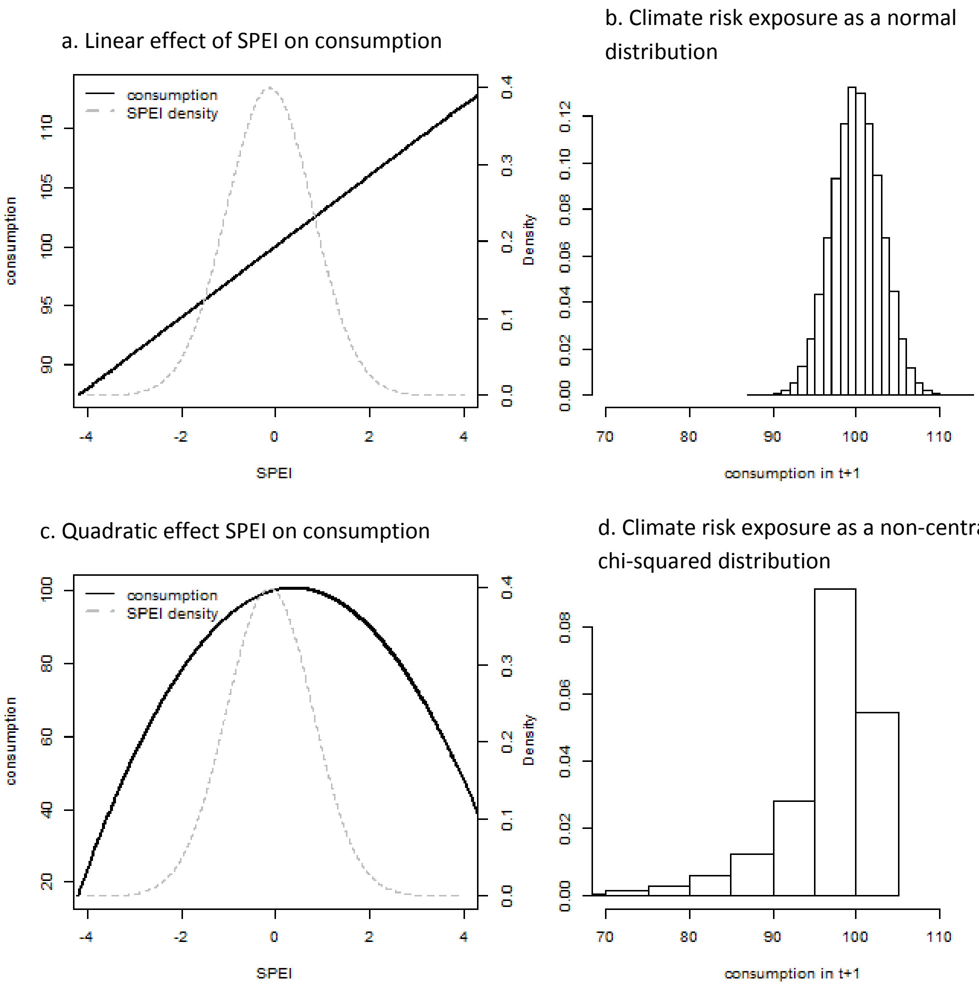
$$g(S, \beta) = \beta_0 + \beta_1 S + \beta_2 S^2 \tag{12}$$

Letting $\beta_0 = 10$, $\beta_1 = 1$ and $\beta_2 = -2$, consumption reach a maximum in conditions slightly moister than normal and decreases both with positive and negative values (plot c of Figure 2.2). Applying the formula in equation (9), the density function of c is:

$$f_c(c) = \sum_{k=1}^{2} \left| 1/\sqrt{{\beta_1}^2 - 4\beta_2(\beta_0 - y)} \right| f_S\left(g_k^{-1}(c)\right) \tag{13}$$

where $g_k^{-1}(c) = \frac{-\beta_1 \pm \sqrt{{\beta_1}^2 - 4\beta_2(\beta_0 - y)}}{2\beta_2}$. We show it in plot d of Figure 2.2.

Figure 2.2: Recovering climate risk exposure



When $g(S)$ is a quadratic function, the probability density function $f_c(c)$ is a non-central chi-squared distribution. Indeed, let us rewrite $g(S)$ as:

$$c = a(S-B)^2 + D \tag{14}$$

where $a = \beta_2$, $B = -\beta_1/2\beta_2$, $D = \beta_0 - (\beta_1/2\beta_2)^2\beta_2$, $S - B \sim N\left(\frac{\beta_1}{2\beta_2}, 1\right)$ and $(S-B)^2$ follows a noncentral chi-squared distribution with 1 of degree of freedom and noncentrality parameter $\lambda = (-B)^2$ (Casella and Berger 2002), i.e. $(S-B)^2 \sim \chi_1^2(\lambda)$.

Substituting back the original β parameters and taking into account the fact that c is a linear transformation of a $\chi_1^2(\lambda)$ distributed variable, the first four central moments of $f_c(c)$ are given by:

$$\mu_1(c) = \beta_2\left(1 + \left(\frac{\beta_1}{2\beta_2}\right)^2\right) + \beta_0 - \left(-\left(\frac{\beta_1}{2\beta_2}\right)^2\right)\beta_2 = \beta_0 + \beta_2 \tag{15}$$

$$\mu_2(c) = 2\beta_2^2\left(1 + 2\left(\frac{\beta_1}{2\beta_2}\right)^2\right) = 2\beta_2^2 + \beta_1^2 \tag{16}$$

$$\mu_3(c) = \beta_2^3\left(8\left(1+3\left(\frac{\beta_1}{2\beta_2}\right)^2\right)\right) = 8\beta_2^3 + 6\beta_2\beta_1^2 \tag{17}$$

$$\begin{aligned}\mu_4(c) &= \beta_2^4\left[12\left(1+2\left(\frac{\beta_1}{2\beta_2}\right)^2\right)^2 + 48\left(1+4\left(\frac{\beta_1}{2\beta_2}\right)^2\right)\right] \\ &= 48 + 12\beta_2^4 + 6\beta_1^4 + 96\beta_1^2\beta_2^2.\end{aligned} \tag{18}$$

While $\beta_0$has only an impact on the mean, $\mu_1(c)$, β_1 and β_2 have an impact on all the central moments of $f_c(c)$.

The advantage of the approach proposed above is hence to offer a direct estimate of the central moments of the climate risk exposure function via the β parameters. It circumvents the use of the residuals as done in traditional model of production risk analysis (John M. Antle 1983).

Lastly, we can specify consumption not only as a function of SPEI, but also as a function of other determinants such as land tenure, agro-ecological zones, development intervention, etc. Furthermore, we can assess the impact of a given variable, say agro-ecological zones, on climate risk exposure by specifying the following function:

$$g(S, G, \beta) = \beta_0 + \beta_1 S + \beta_2 S^2 + \beta_3 G + \beta_4 GS + \beta_5 GS^2 \tag{19}$$

where G is a dummy variable equal to one for household living in a given agro-ecological zone. Solving this quadratic equation, the inverse function of $g(S, G, \beta)$ and its derivative in absolute terms are:

$$g(c)_k^{-1} = \frac{-(\beta_1+\beta_4 G) \pm \sqrt{(\beta_1+\beta_4 G)^2 - 4(\beta_2+\beta_5 G)(\beta_0+\beta_3 G - c)}}{2(\beta_2+\beta_5 G)} \tag{20}$$

and

$$\left|\frac{dg_k^{-1}(c)}{dc}\right| = \frac{1}{\sqrt{(\beta_1+\beta_4 G)^2 - 4(\beta_2+\beta_5 G)(\beta_0+\beta_3 G - c)}}. \tag{21}$$

Equations (20) and (21) are then inserted into equation (9) in order to obtain $f_c(c)$ and we can modify equations (15)-(18) accordingly in order to estimate the central moments in and out of the agro-ecological zone of interest.

We do not do address here any of the classic endogeneity issues: it provides only a framework to estimate various measures of risk under the assumption that the additional variables included in the model, G in the equation above, are exogenous.

2.2.4 Vulnerability Indices

Once the relationship between consumption and climate is established and the properties of climate risk exposure are derived (variance, skewness etc.), we need to summarize this information into a measure with an economic meaning. We use for this a set of vulnerability indices. Our first index is the consumption poverty risk. It has notably been used by Chaudhuri, Jalan, and Suryahadi (2002) and Christiaensen and Subbarao (2000) and it is defined as the probability that a household's consumption falls below the poverty line.

In the framework presented in section 2.2.3, poverty risk is computed by estimating the β parameters of equation (9) or (12) with SPEI and consumption data. Then, we generate with a random number generator a sample of shocks following a standard normal distribution, i.e. an artificial sample of SPEI, and compute a consumption level under each scenario thanks to the β parameters. Poverty risk is then approximated by the percentage of simulated consumption scenarios falling below the poverty line.

In order to illustrate the versatility of the estimation approach introduced in section 2.2.3, we compute the expected poverty shortfall given weather conditions at least as bad as a weather shock of a magnitude expected to occur at most every 25 years.[13] We start by defining a threshold such as an extreme weather event with a 5 years return period, be it a drought or a flood. We then compute the difference between the expected consumption under such conditions or worse the poverty line:

$$V_2(S) = z - \int_{-\infty}^{g(S)} c f_c(c) \mathrm{d}c \tag{22}$$

where z is the poverty line and S is defined in terms of the return period of the event in question. For instance, for drought expected to occur every 25 years, S equals -1.9. The conditional expected poverty shortfall, $V_2(S)$, will hence measure the average cash transfer required to bringing back a household to the poverty line in the case of a weather event with a magnitude expected to recur every 25 years at most.

Lastly, we can compute the relative risk premium with:

$$RRP \approx \left(\frac{AP}{2}\mu_2 - \frac{DS}{6}\mu_3 + \frac{FT}{24}\mu_4\right)/\mu_1 \tag{23}$$

[13] It could be summarized as the combination of the expected poverty shortfall commonly used in vulnerability analysis and the conditional values-at-risk measure used in the finance literature (e.g. Engle and Manganelli 2004). We don't attempt here at providing any rigourous derivation of this index, it is only used for illustrative purposes. Interested readers are referred to Foster et al. (2010) for a review of the use of the Foster-Greer-Thorbecke class of poverty measures with illustration of recent vulnerability analysis, or Hoddinott and Quisumbing (2003)e.g. , e.g. for a more detailed review on vulnerability measurement.

where μ_i are the i[th] central moment as expressed in equations (15)-(18), AP is the coefficient of absolute risk aversion (Pratt) for mean-preserving spread aversion, DS is the coefficient of downside risk aversion (Menezes et al. 1980), for mean-spread-preserving skewness preferences and FT is the coefficient of kurtosis aversion (Rubinstein et al. 2006) for mean-spread-skewness preserving kurtosis aversion. We specify the utility function as follows:

$$U(x) = \frac{U(x)^{1-\gamma}}{1-\gamma} \tag{24}$$

We will conduct some sensitivity analysis on the γ parameter as its values varies according to academic fields and authors (e.g. Holt and Laury 2002; Ligon and Schechter 2003; Yesuf and Bluffstone 2009).

These three vulnerability indices give different perspectives on the climate risk exposure of households. The poverty risk is intuitive, but does not take into account the expected depth of poverty. The conditional expected poverty shortfall index captures downside risk and could be useful for contingency planning[14]. Lastly, the relative risk premium emphasizes the trade-off between expected profit and risk and could be used for targeting the roll-out of private agricultural insurance policies such as weather index insurance. Indeed, the relative risk premium, also known as the implicit cost of risk bearing, is an estimate of household willingness to pay for risk reduction.

2.2.5 Estimation strategy

In order obtain estimates of the β parameters in equation (12), we will use ordinary least squares on the following regression line:

$$c = \beta_0 + \beta_1 S + \beta_2 S^2 + \varepsilon \tag{25}$$

where $\beta_i$are the parameter to estimate, S is the SPEI values of the peak rainfall month of the preceding season and ε is an error term. We will also investigate how weather sensitivity and climate risk exposure varies between as agro-ecological zones with the following regression line:

$$c = \beta_0 + (\beta_1 S + \beta_2 S^2 + 1)\boldsymbol{AEZ\beta_{AEZ}} + \varepsilon \tag{26}$$

where $\boldsymbol{AEZ}$ is a set of dummy variables for the agro-ecological zones and $\boldsymbol{\beta_{AEZ}}$ a vector of parameter to be estimated. Lastly, we will test the impact of access to basic service on vulnerability with the following specification:

[14] A much more thorough examination of its properties would be required however before using it for any policy purpose.

$$c = \beta_0 + (\beta_1 S + \beta_2 S^2) + \boldsymbol{AEZ}\boldsymbol{\beta_{AEZ}} + \boldsymbol{X\beta} + \varepsilon \tag{27}$$

where $\boldsymbol{X}$ is a set of basic services we will describe in the data section.

As the climate variable is exogenous and varies randomly over time, the regression should not suffer from reverse causality bias. In order to control for omitted variable bias, we use household fixed effects in order to control for unobserved heterogeneity between household. In each case, we will present results obtained both with pooled regression and panel fixed effects models.

Given the serial correlation of the error term coming from the use of fixed effects as well as the likely heteroskedasticity, we computed heteroskedasitcity and autocorrelation robust standard errors following Arellano (Arellano 1987) with the plm R CRAN package (Croissant and Millo 2008). Furthermore, all explicative variables are measured at the peasant association (PA) level. It is therefore likely that the error exhibit a certain degree of clustering at the PA level despite the use of household fixed effects. We therefore apply a degree of freedom correction to the variance-covariance matrix of the parameter estimates β_i, $i = 1, \dots, N$. The resulting variance-covariance matrix is hence computed as:

$$V(\widehat{\boldsymbol{\beta}}) = \frac{G}{G-1}\frac{N-1}{N-K}(X^T X)^{-1}\sum_{i=1}^{N} X_i^T \varepsilon_i \varepsilon_i^T X_i \,(X^T X)^{-1} \tag{28}$$

where G is the number of PA, N is the sample size and K is the model rank.

Note that the parameter estimates on SPEI have to be read in terms of standard deviation: a net balance of water one standard deviation away from normal causes a change of β consumption units. The fact that the SPEI is standardized implies that the water balance is measured in terms of local frequencies. This help sorts another source of unobserved heterogeneity: typically, one can assume that a given level of net balance of water is going to have a heterogeneous impact across agro-ecological zones. The standardization implies that we are comparing the net balances of water in terms of their local frequency so that passing from 0 to 1 on the SPEI scale means the same across the country, i.e. a one standard deviation compared to normal conditions.

Lastly, note that the estimated climate risk exposure via OLS is valid for the average household in the sample. Instead of focusing on expected consumption, we can look at the consumption at other quantiles of the consumption sample distribution. It is likely that poorer farmers exhibit higher climate risk exposure because of a lack of *ex ante* and *ex post* risk mitigating options such as irrigated plots, liquid assets (e.g. bullocks and gold ornaments), off-farm jobs, savings and affluent social networks (e.g. relatives working in a nearby town). We can therefore expand the analysis from the climate risk exposure of the average household to the climate risk exposure at different quantiles:

$$f_{c_\tau}(c_\tau) = \sum_{k=1}^{n(c_\tau)} \left| \frac{d g_\tau^{-1}(c_\tau)}{d c_\tau} \right| f_S\big(g_{k\,\tau}^{-1}(c_\tau)\big) \tag{29}$$

where $c_\tau = Q_\tau(c_{it}|S)$ is the conditional quantile of consumption as a function of SPEI.

Panel econometrics methods for quantile regression have been developed by Koenker (2004) and Abrevia and Dahl (2008). They have recently been applied by Bache et. al (2013) to study the impact of prenatal maternal smoking on the dispersion of birthweights and by Dahl et al. (2013) to study the impact of the decentralization of wage bargaining on wage dispersion. As in the classical mean regression panel methods, they allow for the control of unobserved heterogeneity within the sample. The standard errors were automatically computed with bootstraps thanks to the *rqpd* package on R (Koenker and Holst Bache 2014).

It is interesting to note that there has been some confusion between risk and inequality in the literature using quantile regressions. A clear example of the ambiguity surrounding quantile regressions' estimates is the twin papers of Peirera and Martins (2002, 2004) on the impact of education on wages. In a first version of the paper published in *Economics Letters* in 2002, the authors apply quantile regressions at each decile of the wage distribution with education as an explicative variable. Their goal is to estimate the impact of education on wage uncertainty across sixteen European countries. They interpret their results as follows: '[I]f there is a large difference in the estimated coefficients between the first and last decile, meaning that the return is much higher at the upper than at the lower decile, the individual faces a high risk, as the individual can end up at the lower decile. If the difference is small, there is almost no risk' (Telhado Pereira and Silva Martins 2002). Other studies based on the risk interpretation of quantile estimates have followed, both in the banking sector and the literature examining the impact of education on wage.

A second version of the paper, with exactly the same set of data, econometric analysis, results and published by the same authors one year later in *Labour Economics*, is entitled 'Does education reduce wage inequality?'. In the latter paper, the authors give the inequality interpretation of quantile regressions, i.e. a positive difference between higher and lower quantile estimates implies that education increases inequality: '[their] findings imply that schooling may have a positive impact upon within-group wage inequality, as the spread of returns increases for higher educational levels' (Martins and Pereira 2004).[15]

In support of this interpretation, we observe that the law of iterated expectation does not apply in the quantile world. Therefore, while the conditional expectation can be interpreted as the expectation for an average household, the conditional quantile cannot be interpreted as the quantile of the conditional distribution of an average as household. The classic interpretation of quantile regressions' parameters as indicative of sample inequality seems hence more appropriate. We will assume below that SPEI has a rank preserving effect: the rank of each household is not affected by a change in the SPEI value. We can hence interpret the results of a quantile regression estimated at the median for instance as the effect of SPEI on the median household.

To sum-up, we predict expected consumption thanks to a regression of consumption on the SPEI and, optionally, on a set of controls. As the SPEI is distributed according to a standard normal distribution, we can compute with a change of variable the probability density function of consumption, i.e. the climate risk exposure of the average household (or households at other quantile, for instance the poorest quintile). Climate risk exposure and weather sensitivity are then summarized in three indices: poverty risk, expected shortfall and the risk premium.

2.3 Data

The Ethiopian Rural Household Survey (ERHS) is probably the longest running household survey available on development economics, conducted from 1989 to 2009 in seven rounds, with a staggeringly low level of attrition (see Dercon and Kirshan, 1998, for the sample frame design). On top of being freely available on the International Food Policy Research Institute (IFPRI) website, it comes with a great amount of documentation and videos on the data collection process and data issues. For this paper, we use the data files on consumption and community level information.

There are large seasonal fluctuations in consumption as documented by Dercon and Krishnan (2000). As the surveys have not been conducted exactly at the same period of the year over the rounds, we

[15] The rationale behind this is that 'the earnings increment associated to schooling is higher for those individuals whose unobservable characteristics place them at the top of the conditional wage distribution' (idem). It is hence akin to the latent effect interpretation of quantile regression: inequality in conditional wage outcomes is the result of differences in innate ability revealed by quantile regression. Note that this interpretation is, in turn (and quite paradoxically), related to a special case of Kanbur's model (1979) , where risk is represented by the ability risk that an entrepreneur faces when starting a business for the first time, i.e. the uncertainty about his own capacity to run it. Other earlier works (e.g. Friedman, 1953) have drawn the link between risk and inequality. It also echoes the concept of 'veil of ignorance' used in thought experiments by political philosophers to apprehend social contracts and redistribution (Rawls, 1971).

follow Dercon et al. (2012) and drop data from rounds 2, 3 and 4. Indeed, rounds 2 and 4's data were collected in most villages just after harvest, when a household's consumption is expected to be at its maximum. Round 3 is removed in order to have an equally spaced panel (1994, 1999, 2004, 2009) and avoid hence inconsistent estimates due to heterogeneous frequency (Dercon et al. 2012). The other rounds have been performed, on average, 6 to 9 months after harvest.

Ethiopia changed in many aspects between 1989 and 2009. The country's population increased from 50 m to 83 m between 1992 and 2009 (FAO statistics). Meanwhile, the share of the rural population is quite stable although we do observe a slow and constant decline from 88% in 1989 to 83% in 2009. Lastly, the road network almost doubled between 1997 and 2007, although the share of paved roads did not follow suit (decreasing from 15% to 13.7%). In constant 2005 US dollars, GDP per capita had been oscillating around $140 until 2003 before experiencing a steep rise, reaching $213 in 2009, i.e. a 52% increase in 6 years for an average GDP growth of 11% (World Development Indicators, The World Bank, 2014). The domestic food price index grew from 1.6 in 1990 to 1.9 in 2009. Hence it is not clear, *a priori*, if the food security of the rural population has increased or not over time.

The poverty head count ratio at USD $1.25 PPP declined from 60% to less than 40% between 1995 and 2005 (the only available period in the World Bank data bank). Although the share of agriculture in the GDP declined from 61% to 47% over the period 1989-2009, cereal yields and production increased greatly. The yield hovered around 1,180 kg/ha until 2004 before reaching 1,650 kg/ha in 2009, while production had started its climb up by the beginning of the 1990s thanks to a large increase in land under cereal production. In the 2000s', the increase in production is due, in equal proportion, to the increases in yield and area farmed (Taffesse et al. 2011). In 2007, 96% of the cultivated land dedicated to the main crops (cereals, pulses, oilseeds, vegetables, roots crops, fruits and cash crops) was still farmed by smallholders and their harvest in the main production season (*Meher)*, represents 93% of the Ethiopian cereal production (Taffesse et al. 2011). It is hence of primary concern to better assess smallholders' exposure to climate shocks.

We used two sets of data for the computation of the SPEI thanks to the R package SPEI (Beguería and Vicente-Serrano 2013) with the Thornthwaite evapotranspiration index. The precipitation data come from the African Rainfall Climatology Version 2 dataset (ARC2, Novella and Thiaw 2013), providing daily estimates at a resolution of 0.1 decimal degree from 1983 to the present, and are based on a combination of gauge and satellite data. The dataset has been developed as a key input of the Famine Early Warning System Network (FEWSNET), one of the main indicators used by international humanitarian agencies to monitor food security. The temperature data comes from the Climate Prediction Center Global Land Surface Air Temperature Analysis (GHCN+CAM, Fan and Van den Dool

2008).[16] They come as monthly mean surface air temperatures at a 0.5 decimal degree resolution over the period from 1948 to the present. One of its recommended uses is precisely the computation of evapotranspiration indices. Both ARC2 and GHCN+CAMS datasets are matched with the ERHS thanks to a ward-level (*kebele*) administrative boundaries shapefile (Ethiopian Statistical Agency, 2007 census).

The *kebele*, or Peasant Associations (PA) in the rural part of the countries, were founded by the Coordinating Committee of the Armed Forces, Police, and Territorial Army of Ethiopia, also known as the Derg, after the fall of Emperor Haile Selassie in 1974. They are the lowest administrative unit. We have chosen as matching coordinates the centre of each PA computed with centroids of Voronoi. Note that the median area of the EHRS PAs is smaller (50 km^2) than the median aRC2 and GHCN+CAMS cells (120 km^2 and 3,025km^2 on average respectively); they hence constitute a matching metric precise enough for the climate data resolution.[17]

There are three main weather regimes in Ethiopia: the northern part has a bi-modal regime with a long rainy season from June to September and a short rainy season from March to May (regime A); the western part of the country has a mono-modal regime with rainfall from June to September (regime B); and the southern and eastern part has a mono-modal weather regime with rains from February to May (regime C) (NMSA 1996, cited in Abebe 2010). The approximate hand-drawn partition of the country between weather regimes, according to a map of the Ethiopian National Meteorological Agency (1996) reproduced in Abebe (2010), is mapped with long dashed lines in Figure 2.3 (a). Note that according to the ARC2 rainfall data for each PA, the partition is slightly different (dotted line).[18]

Figure 2.3: Elevation and weather regimes (a) and annual precipitation (b)

[16] GHCN Gridded V2 data provided by the NOAA/OAR/ESRL PSD, Boulder, Colorado, USA, from their Web site at http://www.esrl.noaa.gov/psd/.

[17] Area weighted precipitation and temperature means would also have been an option for PAs at the junction of multiple cells, but given the spatial definition of the climate datasets, it would not have affected the results much.

[18] Although the ARC2 dataset would allow estimating the boundaries between weather regimes with more precision, it is outside of the scope of the present paper.

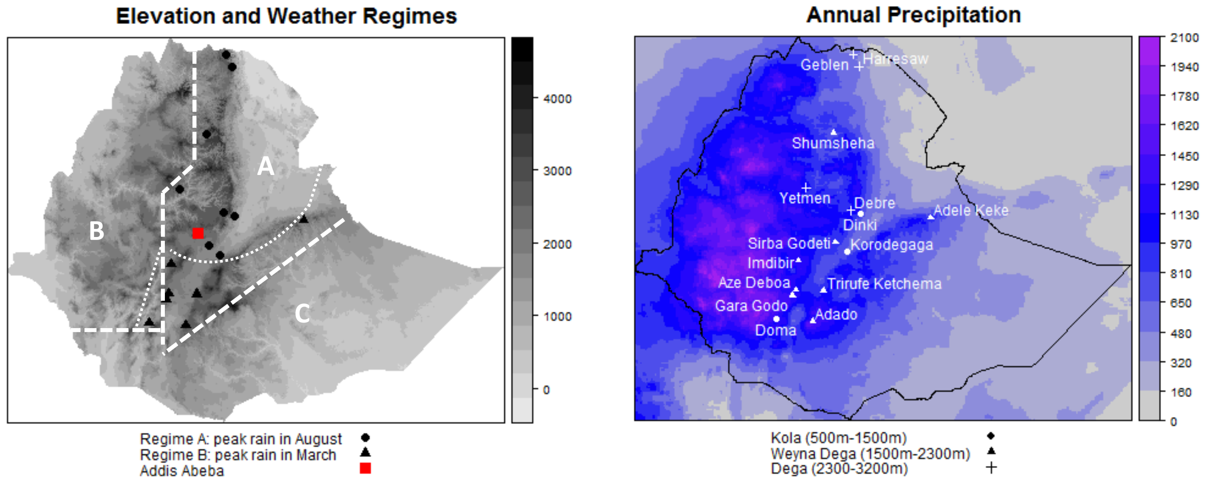
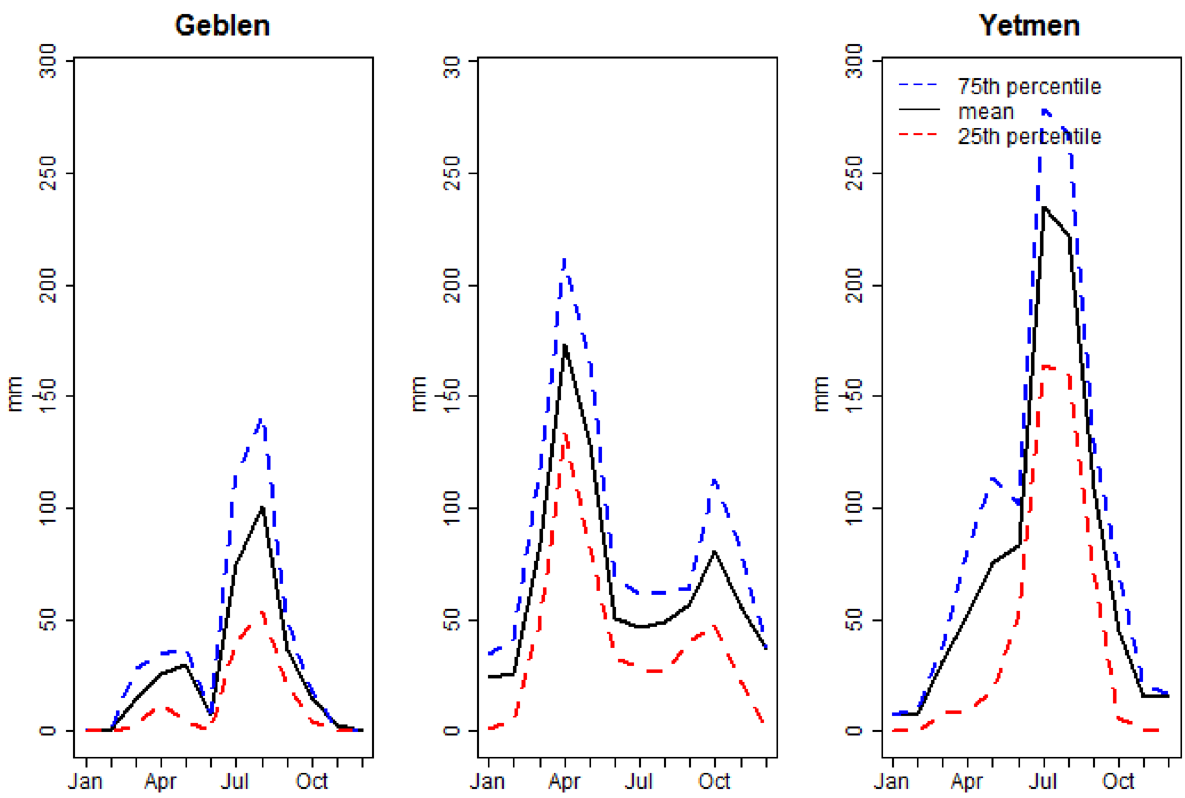


Figure 2.3 a and b: On the left figure, the long dashed lines are the approximate partition of the country between weather regimes according to a 1996 map of reproduced in Abebe (2010) and the dotted lines represent an alternative partition matching the ARC2 data at Pas' locations. The map on the right is the average annual precipitation over the period 1990 to 2013. Although it is clear that precipitation is concentrated on reliefs because of convective rain, there are great differences in precipitation between PAs located at similar altitudes: Geblen receives less than 320mm on average while Yetmen, in the same agro-ecological zone, receives twice as much.

Note that the cumulative level of rainfall varies a lot between PAs in regime A (Figure 2.3 b): normal annual precipitation[19] for Geblen and Harresaw (Tigray region, top North) is only 270 mm while it is 680 mm in Yetmen (Amhara, central North). The PAs located in weather regime C have a maximum amount of cumulative rainfall in March while those located in weather regime A have their maximum in August. We plot in Figure 2.3 b the annual precipitation profile for Geblen (regime A), Doma (regime C) and Yetmen (regime A). We use the climate data for the peak months in the analysis.

Figure 2.4: Monthly precipitation

[19] Normal computed on 1994 to 2013, Hoefsloot 2013, LEAP software.



We use as our dependent variable real consumption per capita as provided in the ERHS. The explicative variables are the 3 months smoothed SPEI at peak rainfall month, the agro-ecological zones, the quality of the road leading to the next town, the distance to the nearest bank, the number of extension agents within the PA and the presence of a non-governmental organisation (NGO) in the PA. Summary statistics are presented in Table 2.1.

Table 2.1: Summary statistics

	Mean	Median	Stand. Dev.	Min.	Max.
Real consumption per capita (birr)	77.63	56.79	74.17	0.88	1,109.39
3-SPEI at peak precipitation month	0.22	0.21	0.91	-1.56	2.21
Remote from a bank (22 km)	0.42	0	0.49	0	1
NGO in the PA	0.16	0	0.36	0	1
Extension agent in the PA	0.76	1	0.43	0	1
Road improvement	0.59	1	0.49	0	1

Although the national figures paint a rather positive picture for recent years, micro level evidence from the ERHS warrants some caution. While the poverty rate hovers between 45% and 50% until 1995 in the ERHS sample, it decreases to 30% in the next 3 rounds (1997, 1999, 2004) before rising again, above 50% in 2009 (Dercon et al. 2012). The average consumption is 78 birr (Ethiopian currency) per month (*circa* USD 18) if one focuses on the 1994, 1999, 2004 and 2009 rounds. There are some substantial variations across years: the 1989, 1994 and 1995 average consumption is around 70 birr; the 1997, 1999 and 2004 average consumption increases to 90 birr while 2009 sees a 34% drop in consumption to 60 birr per month. Consumption per capita includes household-produced food and hence is directly impacted by weather conditions. Details of the real consumption per capita calculation can be found in Dercon and Krishnan (1998). We follow Dercon and Krishnan (1998) in setting the poverty line at the income level required to buy 2,400 calories per day, i.e. 50 birr. The vulnerability indices are hence linked to climate related food insecurity.

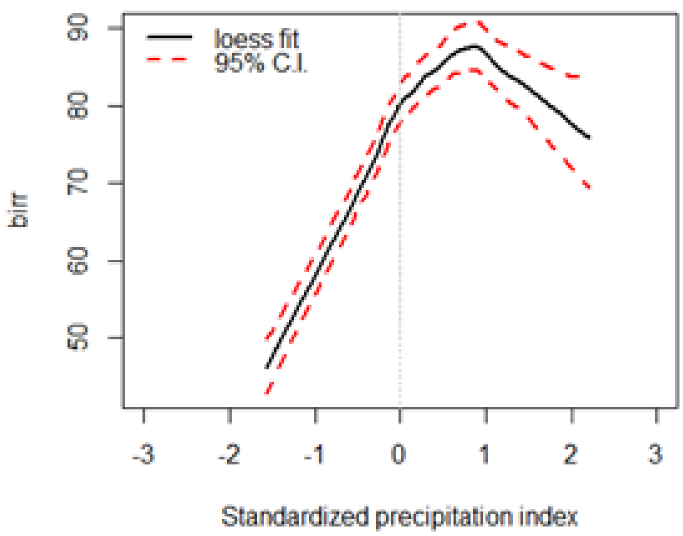
According to the weather regimes identified above, we focus on precipitation in the months of March we are interested in the hydrological conditions affecting agriculture production, we select the three month smoothed SPEI values. We use one year lagged SPEI values as the surveys have been conducted in pre-harvest periods, i.e. when real consumption is still determined by the previous year's harvest. The average SPEI is 0.21, i.e. conditions were on average slightly wetter than normal. The minimum and maximum are respectively -1.57 (2009, in Imdibir) and 2.2 (1994, in Trirufe Ketchema), i.e. dry conditions with a 20 years return periods and wet conditions close to a 100 year return period. Note that consumption prediction conditional on values outside the sample range will have to be treated with caution and can only represent high bound estimates, as it is likely that consumption collapses at higher (lower) SPEI values than the one observed.

The community-level data capture some of the classical development policies. Indeed, road improvement allows better market linkages with the rest of the country and hence offers better marketing opportunities, larger and more stable sets of products for buying, better price smoothing when local production is adversely hit and allows households to enter into new profitable activities (Dercon et al., 2012). Extension agents remain a key development mechanism whereby civil servants are dispatched among rural communities to offer farm management advice and increase the adoption of best farming practices. We express it as a dummy variable equal to one if there is at least one extension agent in the PA. Over time, all PAs got an extension agent. The distance to the nearest bank is also of interest as they are a key channel in providing saving mechanisms, as *ex ante* risk management and credit for adopting more capital intensive inputs. Furthermore, the distance to the nearest bank serves as a proxy of the *remoteness* or *secludedness* of a particular PA as banks are likely to establish branches in local economic centres. We express it as dummy equal to one if the PA is located at more than 22 km from any bank, the latter value being the median sample distance. The presence of an NGO or a development agency might not only have an impact on their sectorial activity, be it education, health or micro-credit; but they can also be an important provider of jobs for the local community. Furthermore, in case of an adverse climatic shock, an NGO might be able to scale up its activity and act as a safety net for the local community. Dercon and Krishnan (2003) showed that food aid provided an insurance mechanism.

2.4 Results

We start by investigating the functional shape of the relationship between real consumption per capita and the standardized precipitation evapotranspiration index (SPEI) with localized polynomial regressions. The smoothing fit is plotted in Figure 2.5 along the 95% confidence intervals computed by performing 1,000 bootstraps with replacements. The relationship is clearly u-shaped, with a maximum at 0.8, i.e. conditions slightly wetter than normal.

Figure 2.5: Influence of weather on consumption per capita



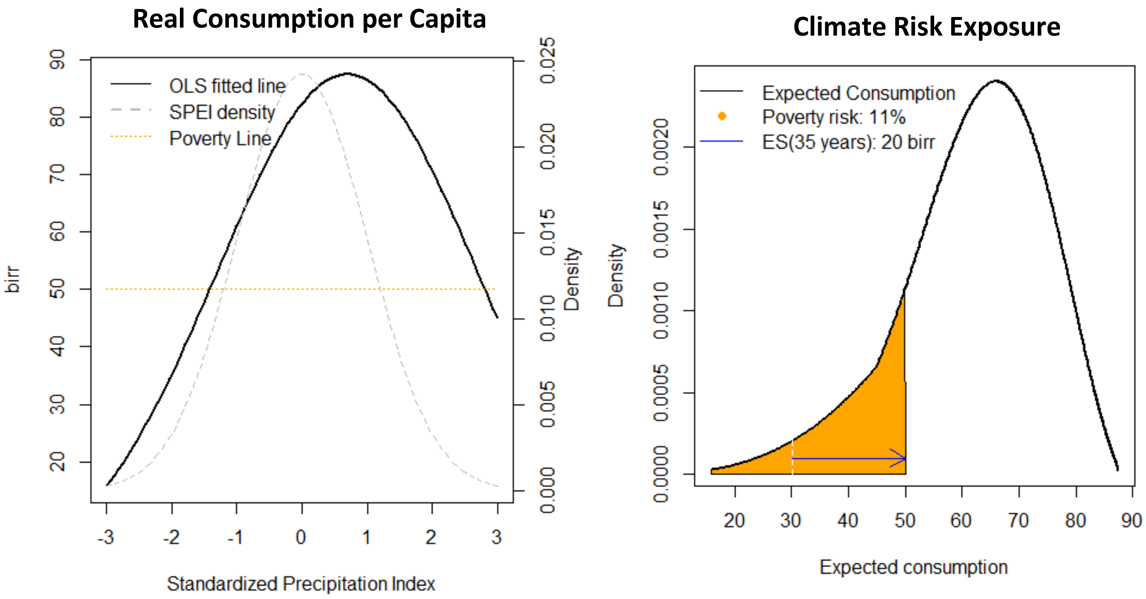
We start the analysis with a simple pooled OLS quadratic regression of consumption on SPEI in order to evaluate the average climate risk exposure in the sample. As the consumption values are very skewed, we apply logarithmic transformation on the consumption variable. The regression we estimate is given by the following expression:

$$\log(c_{it}) = \beta_0 + \beta_1 S_{it} + \beta_2 S_{it}^2 + \varepsilon_{it} \tag{30}$$

where c_{it} is the real consumption per capita of household i at time t, S is the 3-months smoothed SPEI at peak rainfall months and $\beta_j, j = 1,2,3$, are parameters to be estimated. Note that the intercept, β_0, is the log of expected consumption under normal conditions, i.e. when the SPEI equals 0. Results are presented in Table 2.2, column 1.

All parameters are statistically significant (p-value <0.001). The low R^2 should not be a concern as various other factors determine the between variation in the sample distribution of consumption (the size of the land holding, the size of the herd, etc). Nevertheless, a clear pattern emerges from this simple regression: consumption has an inverted U shape in SPEI and reaches its maximum at a SPEI value of 0.7, i.e. in conditions slightly moister than normal, and decreases sharply in drier conditions, crossing the poverty/hunger line at a SPEI value of -1.4, i.e. in severely dry conditions occurring on average every 12 years. Consumption can also fall under the poverty line for extreme precipitation levels, i.e. a SPEI of 2.8 consisting consistent with an extreme flood event. However, such events have only a 0.2% chance of occurring, and hence weight less in farmers' exposure to climate risk. Note, however, that the observed SPEI values in the sample are limited to -1.48 to 2.21, hence predictions outside the sample range have to be considered with care.

Figure 2.6: Real consumption per capita (a) and climate risk exposure (b)



The graph in Figure 2.6 (a) is the fitted consumption line as a function of SPEI. The probability function of the SPEI is superimposed in grey in order to get a better sense of the likelihood of each SPEI value. The area coloured in orange in Figure 2.6 (b) is the probability mass of falling below the hunger line, i.e. 11% in the present case. We also represent the expected shortfall with a 35 years return-period drought (blue arrow, 20 birr). A quick calculation indicates that a 10 year return period drought hitting a region with 100,000 inhabitants would cost a humanitarian agency on average 800,000 birr (*circa* USD 192,000) per month in cash vouchers/transfers to ensure that basic food requirements are met.

Table 2.2: Regressions with agro-ecological zones as additional explicative variables

	Dependent variable: log of real consumption per capita						
		OLS (pooled)		*OLS with fixed effects*	*Quantile regressions with penalized fixed effects*		
	I	*II*	*III*		*1st quartile*	*Median*	*3rd quartile*
SPEI	0.174***		0.125***	0.131***	0.119***	0.137***	0.147***
	(0.017)		(0.017)	(0.017)	(0.02)	(0.017)	(0.02)
SPEI^2	-0.125***		-0.128***	-0.053***	-0.085***	-0.052***	-0.054*
	(0.016)		(0.016)	(0.018)	(0.015)	(0.015)	(0.02)
High altitude ^{D}		0.102***	0.109***		0.046	0.088***	0.126***
		(0.035)	(0.033)		(0.035)	(0.03)	(0.043)
Low altitude ^{D}		-0.156***	-0.200***		-0.207***	-0.111***	-0.083*
		(0.034)	(0.037)		(0.044)	(0.031)	(0.045)
High altitude ^{D} *SPEI			0.222***	0.089***	0.208***	0.152***	0.111***
			(0.029)	(0.031)	(0.036)	(0.03)	(0.035)
Low altitude ^{D} *SPEI			-0.043	0.016	-0.048	-0.081**	-0.05
			(0.037)	(0.054)	(0.056)	(0.038)	(0.04)
High altitude ^{D} *SPEI^2			-0.045*	0.049	-0.005	-0.019	-0.025
			(0.025)	(0.032)	(0.033)	(0.027)	(0.033)
Low altitude ^{D} *SPEI^2			-0.010	-0.076*	-0.013	-0.098***	-0.153***
			(0.035)	(0.044)	(0.046)	(0.038)	(0.045)
Constant	4.103***	4.032***	4.128***	4.043***	3.800***	4.074***	4.391***
	(0.021)	(0.022)	(0.021)	(0.343)	(0.022)	(0.017)	(0.028)
Observations	5,240	5,240	5,240	5,240	5,240	5,240	5,240
R^2	0.045	0.011	0.074	0.056			
Adjusted R^2	0.044	0.011	0.073	0.041			
F Statistic	122.316***	30.077***	52.287***	38.131***			

*Cluster robust standard error in parentheses, *p<0.1; **p<0.05; ***p<0.01, ^{D} stands for dummy variables.*

We then add a series of dummies for the agro-ecological zones, taking the mid-altitude zone (*Weyna-Dega*) as base category, and we interact them with the SPEI variables:

$$\log(c_{it}) = \beta_0 + \beta_1 S_{it} + \beta_2 S_{it}^2 + \beta_3 S_{it} * K_i + \beta_4 S_{it}^2 * K_i + \beta_5 S_{it} * D_i + \beta_6 S_{it}^2 * D_i + \varepsilon_{it} \tag{31}$$

where K stands for the lowlands dummy and D for the highlands dummy. We test for the presence of unobserved heterogeneity with a Lagrange multiplier test (Breusch-Pagan) and a F-test of the model with fixed effects and against pooled OLS (p-value <0.001). The null hypothesis is rejected in all cases with a high confidence level (more than 99.99%); we hence conclude that there are important unobserved effects. We then compare the random effects model against fixed effects models with a

Hausman test and reject the null hypothesis of convergent estimates, preferring the fixed (*within*) effects model. Lastly, we test for the presence of serial correlation threatening the strict exogeneity assumption of the fixed effects model with the Wooldridge test for serial correlation, and fail to reject the null hypothesis of no serial correlation (p-value=0.32). We choose, therefore, a fixed effects model to take into account the households' unobserved heterogeneity. The results are reported in Table 2.2, column 4. For comparison purposes, we present in column 2 and 3 regressions results with pooled OLS when only the altitude dummies are added and where both altitudes and interaction variables are added.

As we see in Table 2.2 column 4, $\hat{\beta}_1$and $\hat{\beta}_2$ decrease compared to Model I, implying that the climate sensitivity in the midlands (*Weyna Dega*) is lower than the average. Furthermore, it appears that the quadratic effect of SPEI is null in the highlands as $\hat{\beta}_6 \approx -\hat{\beta}_2$, i.e. that expected consumption would only increase in SPEI values. This result has to be nevertheless treated with caution given the low level of statistical significance of $\hat{\beta}_6$. By contrast, the lowlands are much more sensitive than the Weyna Dega as $\hat{\beta}_4$ is negative, highly significant and of greater magnitude than $\hat{\beta}_2$.

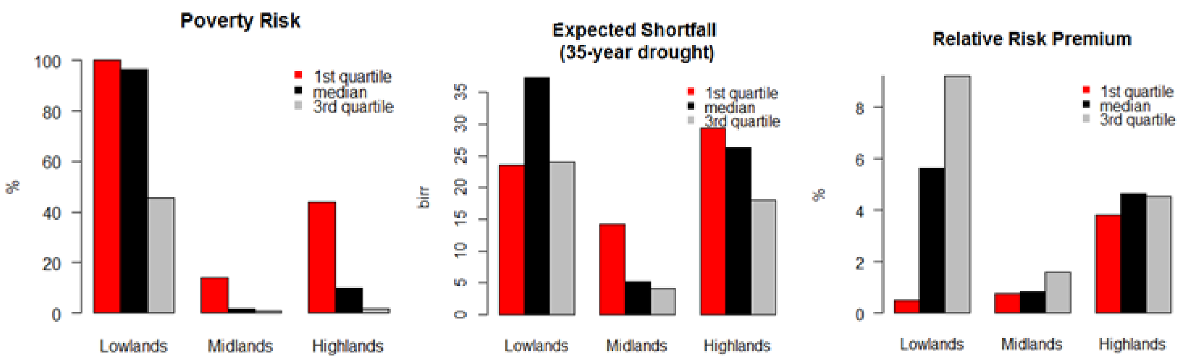
Computing the different indices for each region, the mid-altitude villages have, on average, a poverty risk of 1%, while those located in the highlands of 12% and those in the lowlands of 47%. In terms of expected shortfall, the average household in the midlands is found to be fully resilient even when confronted by a 35 years drought. By contrast, the lowlands have an expected shortfall of 24 birr. These results compare well with Deressa et al. (2009) who also found a greater vulnerability in the lowlands.

We now present the results across a subset of quantiles of the populations estimated with penalized quantile fixed effects quantile regressions (Koenker 2004) and implemented with the package rqpd (Koenker and Holst Bache 2014). The results are reported in Table 2.2, columns 5 to 7. Climate sensitivity does not vary much between agro-ecological zones for the lower quartile in terms of the curvature of consumption. The only significant parameter among the interactions is the interaction of the SPEI expressed in level with the highlands dummy: poor households in high altitude villages reach a maximum consumption in conditions wetter than the rest of the sample. Comparing the interaction terms between the lowlands dummy and the SPEI squared, we see that climate sensitivity increases for households as consumption per capita increases. It suggests, hence, an important trade-off in the lowlands between increase in consumption and decrease in climate sensitivity, the poorer households being stuck in a low risk-low consumption trap, a phenomenon described in the literature on the risk-induced poverty trap.

We present in Figure 2.7 (a), (b) and (c) the 3 vulnerability indices across quantiles and agro-ecological zones.[20] In the lowlands, the lowest quartile is trapped in poverty as its poverty risk is 100%. Furthermore, the median households also face a risk of poverty close to 100% while the 3rd quartile is slightly above 40%. This contrasts with the results found with OLS where the average household had *only* a 47% risk of poverty. Hence, it is likely that the OLS poverty risk estimate was driven downward by the top percentiles of the population. In the midlands and the highlands, the poverty risk is quite low for households above the median although still substantial for the 1st quartile.

The results in terms of the expected shortfall are presented in Figure 2.7 (b). Although the ranking of agro-ecological zones in terms of risk is respected, the differences are much smaller. Furthermore, the ranking within zones changes a lot, e.g. in the lowlands the median 35-year drought expected shortfall is higher than the lower quartile one. The relative risk premium (Figure 2.7 (c)), i.e. the implicit cost of risk computed for a coefficient of relative risk aversion equal to 2, confirms the interpretation of a risk-induced poverty trap by showing that poor households have a smaller relative risk exposure: they have already reduced risk exposure to its maximum at the cost of a decrease in consumption.

Figure 2.7: Poverty risk (a), expected shortfall (35-year drought) (b), Relative risk premium (c)



As the value of the relative risk premium is largely determined by the coefficient of relative risk aversion, we performed a sensitivity analysis assuming that the coefficient of relative risk aversion follows a gamma distribution (see Figure 2.8 a). The distribution was chosen to reflect findings of Yesuf and Bluffstone (2009) based on experimental evidences in Ethiopia. As Holt and Laury (2002) found sensibly lower estimates, we allow for risk preferences below the minimum found in Yesuf and

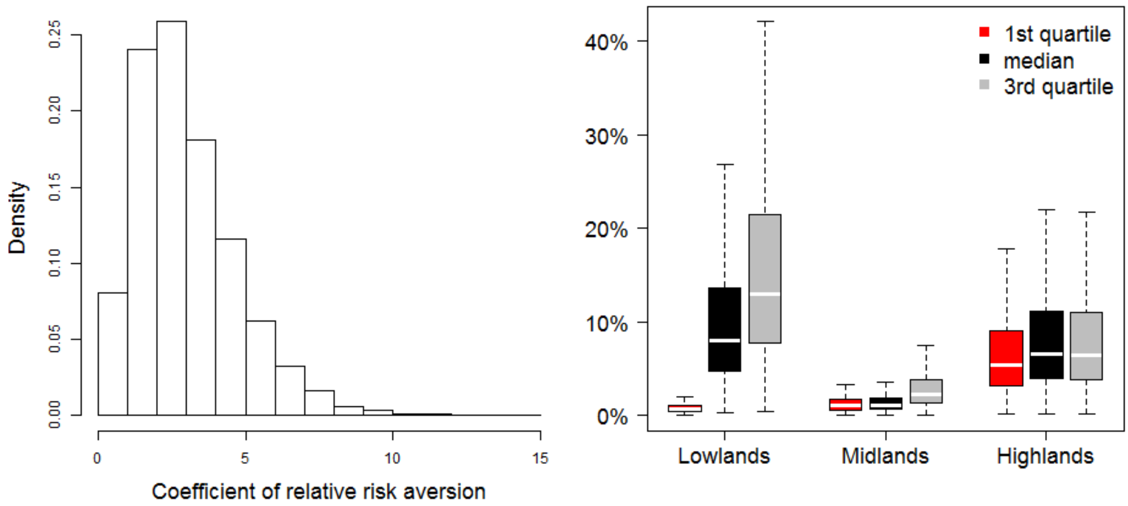
[20] Note that the quantile regressions were run in level to compute the indices because it is *a priori* not clear how to deal with the residuals of exponential quantile regressions when computing the conditional quantiles.

Bluffstone (2009), i.e. between 0 and 1. We present in Figure 2.8 b the resulting box-plot of relative risk premia.

Figure 2.8: Sensitivity analysis on relative risk premia estimated in the quantile regressions

a. Distribution of the coefficient of relative risk aversion used in the sensitivity analysis

b. Relative risk premia according to the risk aversion distribution on plot a.



A policy maker interested in having the greatest impact on average poverty *via*, for instance, the provision of subsidized fertilizers, should look at the poverty risk indicator and target the lowlands. Interestingly, the expected shortfall shows that in the case of a serious drought, it might not be the poorest quartile of the population which will require most help in the lowlands but instead the median households because the latter are more exposed to downside climate shocks. Lastly, the relative risk premium shows that the implicit cost of risk bearing is the highest among richer households, particularly in the lowlands. Hence, the higher quantile of the population manages to get higher consumption at the cost of a large increase in risk and should therefore be willing to swap part of this risk against some kind of consumption insurance, be it index based or of the traditional agricultural kind.

Community level characteristics are only available for rounds 4, 6 and 7, i.e. 1997, 2004 and 2009. As noted in the data section, the 1997 round was conducted earlier in the season and hence might introduce some unobserved heterogeneity. We attempt to control for it by adding a year dummy for 1997. We focus on the presence of an improvement in the road leading to the next town, the number of extension agents within the PA and the distance to the nearest bank, and the presence of a non-governmental and/or international organization office in the PA. The results are presented in Table 2.3.

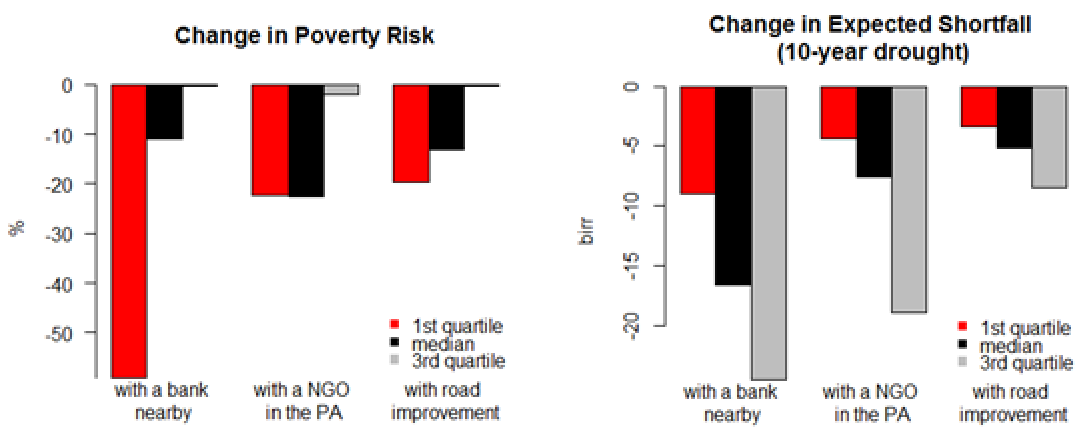
Table 2.3: Regressions with community development factors as additional explicative variables

	Dependent variable: log of real consumption per capita			
	panel Linear	*Quantile regressions with penalized fixed effects*		
	(1)	1st quartile	Median	3rd quartile
SPEI	0.133^{***}	0.188^{***}	0.148^{***}	0.187^{***}
	(0.020)	(0.025)	(0.020)	(0.024)
SPEI^2	-0.087^{***}	-0.130^{***}	-0.130^{***}	-0.124^{***}
	(0.016)	(0.017)	(0.016)	(0.019)
1997 ^{D}	0.163^{***}	0.137^{***}	0.207^{***}	0.139^{***}
	(0.033)	(0.052)	(0.038)	(0.049)
High altitude ^{D}		0.086^{**}	0.045	0.082^{**}
		(0.038)	(0.030)	(0.033)
Low altitude ^{D}		-0.076^{**}	-0.106^{***}	-0.086^{**}
		(0.036)	(0.035)	(0.039)
Bank ^{D}	-0.232^{***}	-0.284^{***}	-0.341^{***}	-0.324^{***}
	(0.041)	(0.039)	(0.036)	(0.042)
NGO ^{D}	0.288^{***}	0.101^{**}	0.126^{***}	0.191^{***}
	(0.037)	(0.049)	(0.045)	(0.046)
Extension agents	0.092^{***}	0.113^{***}	0.115^{***}	0.065^{**}
	(0.024)	(0.029)	(0.026)	(0.031)
Road improvement ^{D}	0.054^{*}	0.044	0.090^{***}	0.150^{***}
	(0.029)	(0.031)	(0.030)	(0.034)
Constant	4.014^{***}	3.762^{***}	4.082^{***}	4.398^{***}
	(0.385)	(0.042)	(0.041)	(0.046)
Observations	3,892	3,892	3,892	3,892
R^2	0.108			
Adjusted R^2	0.071			
F Statistic	44.150^{***}			

Cluster robust standard error in parentheses, $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$, ^{D} stands for dummy variables.

The 1997 dummy is positive, as expected, because the 1997 round was conducted earlier in the season when consumption is higher. The distance to the bank dummy, equal to one if the PA is located at more than 22 km from any bank, is strongly negative: the average household in such a PA has an expected consumption per capita 21% lower than those in PAs closer to a bank. As mentioned in the data section, the presence of a bank might signal that the local economy is particularly dynamic or wealthy. It is hence not clear *a priori* if the positive coefficient on the bank distance dummy does not results from endogeneity caused by the effect of an unobserved variable left in the error term. Note that the effect is quite stable across quantiles of the population (although lower). By contrast, the presence of an NGO office in the PA benefits mostly the median household and above. This might be linked to the fact that jobs created by NGOs tend to benefit the better educated and wealthier households, or it might reveal the difficulty for NGOs to reach the poorest of the poor. Interestingly, road improvement seems again to be of greatest benefit to richer households as no consumption-increasing effect linked to road improvement is found significant in the 1st quartile regression. The number of extension agents has a positive and significant effect: each additional extension agent increases expected consumption by 10% and has the most impact on the median of the distribution. Note that the top quartile does benefit the least from the presence of an extension agent. A possible explanation is that households at the top of the consumption distribution already know and apply best and recommended farming techniques.

Figure 2.9: Change in poverty risk (a) and Change in expected shortfall (10-year drought) (b)



We present in Figure 2.9 (a) and (b) the impact on poverty risk and on the expected shortfall (computed for a 10 years drought) of the variables found significant at different quantiles of the population. As expected, the greatest effect is clearly the bank dummy. It is likely that the bank captures the effect of living close to a dynamic economic centre where banks choose to open a new branch rather than the effect of the bank itself. This large effect might be due to more off-farm opportunities or to the support from relatives living in these economic centres. Note that, however, in the case of a 10-year drought, poor households living near economic centres are not much less exposed than their counterparts in more remote regions, as shown with the expected shortfall. By contrast, richer categories are much more exposed in remote parts compared to their counterparts living close to an economic centre. NGOs and road improvements have a similar effect on the risk of poverty and expected shortfall. Both are, incidentally, positively correlated and it is likely that logistical reasons favour the installation of NGOs in PAs with better road access. Again, we see this mirror relationship between poverty risk and expected shortfall: these are the poorest households who benefit the most in terms of poverty risk reduction but the richest ones in terms of reduction of downside risk.

2.5 Conclusion

This paper introduces a new framework to estimate climate risk exposure at the household level with the SPEI as its building block. It is based on the combination of climate data and household microeconomic data. The main advantage of this approach is that it is based on locally and frequency based weather scenarios allowing different measures of climate vulnerability. Furthermore, as the SPEI is computed over several decades, it properly captures climate risk exposure rather than the short-term, running-season production risk exposure estimated with classic microeconometric methods of production risk estimation. A limitation of the proposed methodology is that it is quite demanding in terms of the large dataset required. Indeed, the estimation of the climate risk exposure relies on the assumption of observing a large range of SPEI values in the sample either thanks to a long panel or thanks to a large geographical spread. We note, however, that the number of microeconomic panel datasets keeps increasing so that this limitation is likely to fade in coming years.

Another advantage of this approach is that it is quite simple and hence is able to accommodate quantile regressions. Instead of being forced to think about the average household, one can broaden the analysis to other parts of the sample distribution. Several indices are proposed to summarize climate risk exposure. The most actionable from a policy standpoint is likely to be the expected shortfall, also known as the conditional value-at-risk.

We illustrate the methodology with a case study on Ethiopia using the Ethiopian Rural Household Survey (ERHS) and we combine it with SPEI values estimated with the African Rainfall Climatology Version 2 dataset and Climate Prediction Center Global Land Surface Air Temperature Analysis (GHCN+CAMS). Results show that the PAs located in the Kolla agro-ecological zone are the most exposed to climate. The results are in line with Deressa et al. (2009), although we do find greater differences between agro-ecological zones. Furthermore, we find that while poor households in the most remote PAs are almost as resilient to 10-year return period droughts as poor households living in the vicinity of towns (within 20 km), the contrary is true for richer households: the ones living in remote parts of Ethiopia are much more at risk than their suburban counterparts.

2.6 References

Abramowitz, Milton and Stegun, Irene A (1972), *Handbook of mathematical functions* (1: Dover New York).

Abrevaya, Jason and Dahl, Christian M (2008), 'The effects of birth inputs on birthweight: evidence from quantile estimation on panel data', *Journal of Business & Economic Statistics,* 26 (4), 379-97.

Antle, John M (1987), 'Econometric estimation of producers' risk attitudes', *American Journal of Agricultural Economics,* 69 (3), 509-22.

Antle, John M. (1983), 'Testing the Stochastic Structure of Production: A Flexible Moment-Based Approach', *Journal of Business & Economic Statistics,* 1 (3), 192-201.

Arellano, Manuel (1987), 'Computing Robust Standard Errors for Within Group Estimators', *Oxford bulletin of Economics and Statistics,* 49 (4), 431-34.

Bache, Stefan Holst Milton, Dahl, Christian Møller, and Kristensen, Johannes Tang (2013), 'Headlights on tobacco road to low birthweight outcomes', *Empirical Economics,* 44 (3), 1593-633.

Beguería, Santiago and Vicente-Serrano, Sergio M. (2013), 'SPEI: Calculation of the Standardised Precipitation-Evapotranspiration Index', (R package version 1.6 edn.).

Casella, George and Berger, Roger L (2002), *Statistical inference* (2: Duxbury Pacific Grove, CA).

Chambers, Robert G. and Quiggin, John (1998), 'Cost Functions and Duality for Stochastic Technologies', *American Journal of Agricultural Economics,* 80 (2), 288-95.

Chaudhuri, Shubham, Jalan, Jyotsna, and Suryahadi, Asep (2002), 'Assessing Household Vulnerability to Poverty from Cross-sectional Data: A Methodology and Estimates from Indonesia', *Discussion Paper Series* (Columbia University).

Christiaensen, Luc J and Boisvert, Richard N (2000), *On measuring household food vulnerability: Case evidence from Northern Mali* (Department of Agricultural, Resource, and Managerial Economics, Cornell University).

Croissant, Yves and Millo, Giovanni (2008), 'Panel Data Econometrics in R: The plm Package', *Journal of Statistical Software,* 27 (2).

Dahl, Christian M, Le Maire, Daniel, and Munch, Jakob R (2013), 'Wage dispersion and decentralization of wage bargaining', *Journal of Labor Economics,* 31 (3), 501-33.

Dell, Melissa, Jones, Benjamin F., and Olken, Benjamin A. (2014), 'What Do We Learn from the Weather? The New Climate-Economy Literature', *Journal of Economic Literature,* 52 (3), 740-98.

Dercon, Stefan and Krishnan, Pramila (1998), *Changes in poverty in rural Ethiopia 1989-1995: measurement, robustness tests and decomposition* (Centre for the Study of African Economies, Institute of Economics and Statistics, University of Oxford).

--- (2000), 'Vulnerability, seasonality and poverty in Ethiopia', *The Journal of Development Studies,* 36 (6), 25-53.

--- (2003), 'Food aid and informal insurance', *WIDER Discussion Papers, World Institute for Development Economics,* 09.

Dercon, Stefan and Christiaensen, Luc (2011), 'Consumption risk, technology adoption and poverty traps: Evidence from Ethiopia', *Journal of development economics,* 96 (2), 159-73.

Dercon, Stefan, Hoddinott, John, and Woldehanna, Tassew (2012), 'Growth and chronic poverty: Evidence from rural communities in Ethiopia', *Journal of Development Studies,* 48 (2), 238-53.

Deressa, Temesgen T, Hassan, Rashid M, and Ringler, Claudia (2009), *Assessing household vulnerability to climate change* (Intl Food Policy Res Inst).
Di Falco, Salvatore and Chavas, Jean-Paul (2006), 'Crop genetic diversity, farm productivity and the management of environmental risk in rainfed agriculture', *European review of agricultural economics,* 33 (3), 289-314.
--- (2009), 'On crop biodiversity, risk exposure, and food security in the highlands of Ethiopia', *American Journal of Agricultural Economics,* 91 (3), 599-611.
Engle, Robert F and Manganelli, Simone (2004), 'CAViaR: Conditional autoregressive value at risk by regression quantiles', *Journal of Business & Economic Statistics,* 22 (4), 367-81.
Fan, Yun and Van den Dool, Huug (2008), 'A global monthly land surface air temperature analysis for 1948–present', *Journal of Geophysical Research: Atmospheres (1984–2012),* 113 (D1).
Foster, James, Greer, Joel, and Thorbecke, Erik (2010), 'The Foster–Greer–Thorbecke (FGT) poverty measures: 25 years later', *The Journal of Economic Inequality,* 8 (4), 491-524.
Friedman, Milton (1953), 'Choice, chance, and the personal distribution of income', *The Journal of Political Economy*, 277-90.
Greene, William H (2003), *Econometric analysis* (Pearson Education India).
Groom, Ben, et al. (2008), 'The story of the moment: risk averse cypriot farmers respond to drought management', *Applied Economics,* 40 (3), 315-26.
Guttman, Nathaniel B (1998), 'Comparing the Palmer drought index and the standardized precipitation index', *Journal of the American Water Resources Association,* 34 (1), 113-21.
--- (1999), 'Accepting the standardized precipitation index: a calculation algorithm', *Journal of the American Water Resources Association,* 35 (2), 311-22.
Hardaker, J. B., et al. (2004), *Coping With Risk in Agriculture* (2 ed. edn.; Cambridge: CABI).
Hayes, Michael J, et al. (1999), 'Monitoring the 1996 drought using the standardized precipitation index', *Bulletin of the American Meteorological Society,* 80 (3), 429-38.
Hazell, Peter BR (1992), 'The appropriate role of agricultural insurance in developing countries', *Journal of International Development,* 4 (6), 567-81.
Hidalgo, F Daniel, et al. (2010), 'Economic determinants of land invasions', *The Review of Economics and Statistics,* 92 (3), 505-23.
Hoddinott, John and Quisumbing, Maria Agnes R (2003), *Methods for microeconometric risk and vulnerability assessments* (World Bank, Social Protection).
Holt, Charles A and Laury, Susan K (2002), 'Risk aversion and incentive effects', *American economic review,* 92 (5), 1644-55.
Just, Richard E and Pope, Rulon D (1978), 'Stochastic specification of production functions and economic implications', *Journal of Econometrics,* 7 (1), 67-86.
Kanbur, Steven M (1979), 'Of risk taking and the personal distribution of income', *The Journal of Political Economy*, 769-97.
Kim, Kwansoo, et al. (2014), 'Rice, irrigation and downside risk: a quantile analysis of risk exposure and mitigation on Korean farms', *European Review of Agricultural Economics,* 41 (5), 775-815.
Koenker, Roger (2004), 'Quantile regression for longitudinal data', *Journal of Multivariate Analysis,* 91 (1), 74-89.
Koenker, Roger and Bassett, Gilbert (1978), 'Regression quantiles', *Econometrica: journal of the Econometric Society*, 33-50.

Koenker, Roger and Holst Bache, Stefan (2014), 'rqpd: Regression Quantiles for Panel Data', *R package version 0.6/r10, http://R-Forge.R-project.org/projects/rqpd/*.
Koundouri, Phoebe, et al. (2009), 'The effects of EU agricultural policy changes on farmers' risk attitudes', *Eur Rev Agric Econ*, jbp003.
Kumbhakar, Subal C and Tveterås, Ragnar (2003), 'Risk preferences, production risk and firm heterogeneity*', *The Scandinavian Journal of Economics,* 105 (2), 275-93.
Ligon, Ethan and Schechter, Laura (2003), 'Measuring Vulnerability*', *The Economic Journal,* 113 (486), C95-C102.
Lobell, David B, Schlenker, Wolfram, and Costa-Roberts, Justin (2011a), 'Climate trends and global crop production since 1980', *Science,* 333 (6042), 616-20.
Lobell, David B, et al. (2011b), 'Nonlinear heat effects on African maize as evidenced by historical yield trials', *Nature Climate Change,* 1 (1), 42-45.
Love, H Alan and Buccola, Steven T (1991), 'Joint risk preference-technology estimation with a primal system', *American Journal of Agricultural Economics,* 73 (3), 765-74.
Martins, Pedro S and Pereira, Pedro T (2004), 'Does education reduce wage inequality? Quantile regression evidence from 16 countries', *Labour Economics,* 11 (3), 355-71.
McKee, Thomas B, Doesken, Nolan J, and Kleist, John (1993), 'The relationship of drought frequency and duration to time scales', *Proceedings of the 8th Conference on Applied Climatology* (17: American Meteorological Society Boston, MA), 179-83.
--- (1995), 'Drought monitoring with multiple time scales', *Ninth Conference on Applied Climatology. American Meteorological Society, Boston*.
Menezes, Carmen, Geiss, Charles, and Tressler, John (1980), 'Increasing downside risk', *American Economic Review,* 70 (5), 921-32.
Moss, Charles B (2010), *Risk, uncertainty and the agricultural firm* (World Scientific).
Novella, Nicholas S and Thiaw, Wassila M (2013), 'African rainfall climatology version 2 for famine early warning systems', *Journal of Applied Meteorology and Climatology,* 52 (3), 588-606.
Palmer, Wayne C (1965), *Meteorological drought* (30: US Department of Commerce, Weather Bureau Washington, DC, USA).
Rawls, John (1971), 'A theory of justice Oxford University Press', *New York*.
Rubinstein, Mark, Jurczenko, Emmanuel, and Maillet, Bertrand (2006), *Multi-moment asset allocation and pricing models* (399: John Wiley & Sons).
Saastamoinen, Antti (2013), 'Heteroscedasticity or Production Risk? A Synthetic View', *Journal of Economic Surveys*.
Sandmo, Agnar (1971), 'On the Theory of the Competitive Firm Under Price Uncertainty', *The American Economic Review,* 61 (1), 65-73.
Schlenker, Wolfram and Roberts, Michael J (2009), 'Nonlinear temperature effects indicate severe damages to US crop yields under climate change', *Proceedings of the National Academy of sciences,* 106 (37), 15594-98.
Schlenker, Wolfram and Lobell, David B (2010), 'Robust negative impacts of climate change on African agriculture', *Environmental Research Letters,* 5 (1), 014010.
Smale, Melinda, et al. (1998), 'The contribution of genetic resources and diversity to wheat production in the Punjab of Pakistan', *American Journal of Agricultural Economics,* 80 (3), 482-93.
Taffesse, Alemayehu Seyoum, Dorosh, Paul, and Asrat, Sinafikeh (2011), 'Crop production in Ethiopia: regional patterns and trends', (Ethiopia ESSP II Working Paper No. 0016:

Development Strategy and Governance Division, International Food Policy Research Institute, Ethiopia Strategy Support Program II).
Telhado Pereira, Pedro and Silva Martins, Pedro (2002), 'Is there a return–risk link in education?', *Economics Letters,* 75 (1), 31-37.
Thornthwaite, Charles Warren (1948), 'An approach toward a rational classification of climate', *Geographical review*, 55-94.
Vicente-Serrano, Sergio M, Beguería, Santiago, and López-Moreno, Juan I (2010), 'A multiscalar drought index sensitive to global warming: the standardized precipitation evapotranspiration index', *Journal of Climate,* 23 (7), 1696-718.
Wells, Nathan, Goddard, Steve, and Hayes, Michael J (2004), 'A self-calibrating Palmer drought severity index', *Journal of Climate,* 17 (12).
Yesuf, Mahmud and Bluffstone, Randall A (2009), 'Poverty, risk aversion, and path dependence in low-income countries: Experimental evidence from Ethiopia', *American Journal of Agricultural Economics,* 91 (4), 1022-37.

2.7 Appendix: Thornthwaite Potential Evapotranspiration index

We follow here closely the presentation of the Thornthwaite Potential Evapotranspiration index (1948) given in Vicente-Serrano et al. (2010). The monthly PET in millimetre at a given location is obtained in the following manner:

$$PET = 16K\left(\frac{10T}{I}\right)^{m} \tag{32}$$

where T is the monthly mean temperature, I is a heat index computed as:

$$I = \sum_{j=1}^{12}\left(\frac{T_j}{5}\right)^{1.514} \tag{33}$$

where j=1, …, 12 for the twelve months; the coefficient m is computed as:

$$m = 6.75 * 10^{-7} * I^3 - 7.71 * 10^{-5} * I^2 + 1.79 * 10^{-2} * I + 0.492 \tag{34}$$

and K is a correction coefficient computed as function of the latitude and the month:

$$K = \left(\frac{N}{12}\right)\left(\frac{NDM}{30}\right) \tag{35}$$

where NDM is the number of days in the month and N is the maximum number of sun hours computed as:

$$N = \left(\frac{24}{\pi}\right)w_s \tag{36}$$

where π is the number pi and w_s is the hourly angle of sun rising calculated using:

$$w_s = arccos(-\tan\varphi * \tan\delta) \tag{37}$$

where φ is the latitude in radian and δ is the solar declination in radian, calculated using:

$$\delta = 0.4093\sin\left(\frac{2\pi J}{365} - 1.405\right) \tag{38}$$

Where J is the average Julian day of the month.

3 Disentangling the Benefits from Agricultural Innovations: Evidence from a combination of open and double-blind experiments in Tanzania

Xavier Vollenweider[21], Erwin Bulte[22], Salvatore Di Falco[23] and Menale Kassie[24]

Abstract

We provide an assessment of the importance of the role of farmers' behaviour in driving the increase in yield of improved maize seeds. The study is based on the combination of an open and double-blind randomized controlled trials (RCTs) conducted in 2013 in two regions of Tanzania with 560 farmers. In the open RCT, farmers were told about which types were allocated to them: half of the farmers received improved seeds, the other half local seeds. The same was done in the double-blind RCT except that farmers were told they had one chance out of two of getting the improved seeds. It allows us to disentangle the increase in yield caused by the improved seeds from the increase in yield which depends on a change in farmers' behaviour. Our main empirical contribution is to show that the behavioural response to improved seeds plays a central role in driving the increase in yield. In our experiment, close to 50% of the increase in yield measured in the open RCT would not have materialized without the behavioural response.

[21] Department of Geography and Environment, London School of Economics and Political Science, United Kingdom. Contact author: Xavier Vollenweider: X.Y.Vollenweider@lse.ac.uk.
[22] Development Economics Group, Wageningen University, The Netherlands.
[23] Department of Economics, University of Geneva, Switzerland.
[24] CIMMYT (International Maize and Wheat Improvement Center), Nairobi, Kenya.

3.1 Introduction

Food insecurity is a major problem in sub-Saharan Africa. The productivity of food crops has been low and decreasing for more than a decade. In Tanzania, for instance, maize yields went from 1.9 tons per hectare (ha) in 2000 to 1.2 tons per ha in 2012 (according to FAOSTAT). Slow adoption of new agricultural technologies is widely regarded as a key determinant of the current state of affairs (Doss 2003; Evenson and Gollin 2003). High yield varieties drove the Green Revolution in Asia and could provide increases in agricultural productivity across Africa as well, stimulating economic growth and facilitating the transition from low productivity subsistence agriculture to a productive, agro-industrial economy (World Bank 2008) . Their uptake in Africa, however, is still limited and far from complete (Foster and Rosenzweig 2010). The literature has provided a plethora of explanations, including lack of access to information, inputs, credit, or risk preferences (e.g. Diagne and Demont 2007). Recently, Suri (2011) used panel data and a random coefficient model to show that the adoption of hybrid maize is simply not profitable for wide swaths of the farming population. If innovations are not sufficiently adapted to the needs and requirements of local farmers (Doss 2003), individuals behave perfectly rationally when deciding not to adopt.

In this paper we use a set of field experiments to address the debated issue of the impact of improved seeds by probing the benefits associated with the adoption of an open pollinated maize variety in Tanzania. Maize is the staple food in the region, and improving productivity in maize is likely to be a key impetus for improving local rural development and food security. Assessing the benefits of adoption is complicated because of two reasons. First, gains likely vary across farmers, depending on geophysical conditions, farm size, and other characteristics affecting output. Heterogeneous treatment effects associated with adoption have been analysed in detail for the case of hybrid maize in Kenya by Suri (2011). Second, subjects may adjust their behaviour following adoption (Bulte et al. 2014; List 2011). For example, changes in the quality of seeds may invite farmers to allocate more or less fertiliser or effort to cultivation.

To examine the importance of the behavioural response as compared to the pure genetic improvement effect, we combine data from two field experiments organised in two regions in Tanzania. In both experiments we provide improved maize to groups of farmers. The first experiment is a conventional (or open) randomised experiment (randomised control trial, RCT). Farmers are randomly allocated either to a treatment group receiving the improved maize or to a control group receiving a traditional variety. Both groups are told which type of seeds they received. The second experiment is a double-blind RCT where neither the participating farmers nor the enumerators are informed about the allocation of the two types of seeds. By comparing the outcomes across the two experiments we can distinguish between the genetic improvement and behavioural effects, and obtain a better understanding of the sources of the innovation's impact on yield than either experiment could produce alone. In addition to obtaining an improved assessment of the impact on yield of this key innovation, we disaggregate the total effect of adoption into a pure genetic improvement effect (i.e., the increase in yields that may be attributed to superior seeds) and the increase in yields that depends on the behavioural response of the farmer to the improved seeds. Indeed, as the information about the treatment status is randomized thanks to the combination of an open and double-blind RCT, related changes in behaviour as well as their impact on yield can be estimated.

When assessing the effect of an intervention, one can focus either on its total effect or on its net effect, i.e. either on the total derivative or on the partial derivative of the production function that the intervention seeks to optimize (Glewwe et al. 2004). Policy makers tend to focus on the total effect of the intervention. The behavioural change of participants is factored in as a benefit of the intervention, even if it represents a cost for them. An alternative approach is to focus on the net effect of the intervention, i.e. net of the additional cost of effort provided by the participants. If the behavioural responses are not observed, then the estimates of the *net* impact of the intervention will be biased (Chassang et al. 2012; Duflo et al. 2008) as the estimation will provide an estimate of the *total* effect of the intervention. In the case of improved seeds, failing to account for the opportunity costs of labour, or any other complementary inputs, implies that the *net* benefits of adoption are over- or underestimated (depending on whether the behavioural response implies the supply of complementary inputs in greater or smaller quantities, respectively). In contrast, if the behavioural response is observed, and fully controlled for in a regression framework, then the indirect effects of adoption via changes in complementary inputs are not attributed to the adoption. Further complicating matters, the behavioural response may be overly optimistic or pessimistic. Farmers may easily overshoot or under-supply complementary inputs in the short run, biasing initial assessments of (potential) profitability. The crux for distinguishing the total impact from the net impact is to get an estimate of the contribution of the behavioural change. The combination of an open and a double-blind RCT provides such an estimate.

The use of double blind procedures is still rare in field experiments. To our knowledge, an exception is Bulte et al. (2014), who study the productivity effect of improved cowpea seeds.[25] Unfortunately, their study is affected by significant attrition and small sample size. Seeds were offered to farmers close to the planting date, and some farmers had already finalized their cropping plans for the season.

[25] Other notable exceptions are found in health economics – see the work on the Work and Iron Status Evaluation (WISE) conducted by Duncan Thomas in Indonesia and Boisson et al. (2010) on water filtration devices in the Democratic Republic of Congo.

Moreover, cowpea is a secondary crop in the study region, and some farmer chose not to grow it. We build and improve upon Bulte et al. (2014) in different ways in an effort to reduce attrition: we distributed seeds earlier in the season, and we used a crop that is the key food staple in the region (maize). As a result our attrition rate is fairly low (9%). We also up-scaled the size of the experiment, and made some changes to the design (details are provided in section 3.4).[26]

The paper is organized as follows: in section 3.2, we provide an overview of the theory behind our methodology. In section 3.3, we present our identification strategy. In section 3.4, after a brief introduction to the study regions, we describe the two experiments and the data. Section 3.5 presents the results and section 3.6 concludes.

3.2 Background

We start by briefly reviewing how a randomized controlled trial (RCT) can provide an estimate of the effect of improved seeds. We then distinguish between the effect caused by the improved seeds alone from the effect which depends on the behavioural response to the improved seeds, for instance an increase in complementary inputs such as fertilisers or labour.

Say we want to estimate the effect of improved seeds on output with a conventional RCT. Farmers are randomly allocated to either of two groups: the first one receives improved seeds, the second one local seeds.

Following the model of Rubin (1974), the potential output of farmer i, Y_i, can be written as:

$$Y_i = \begin{cases} Y_{1i} & \text{if } \tau_i = 1 \\ Y_{0i} & \text{if } \tau_i = 0 \end{cases} \tag{1}$$

where $\tau_i = 1$ if the farmer receives improved seeds and $\tau_i = 0$ otherwise. This can be expressed as:

$$Y_i = Y_{0i} + (Y_{1i} - Y_{i0})\tau_i \tag{2}$$

Assuming that the effect of the improved seed is constant for all farmers, $Y_{1i} - Y_{0i} = \rho$, equation (2) can be re-expressed in regression form as:

$$Y_i = \alpha + \rho\tau_i + \eta_i \tag{3}$$

[26] Double-blind implies that seeds should not be recognizable. We followed Bulte et al. (2014) by treating the improved and local seeds with a reddish food colorant in the double-blind RCT. We modified their design, however, by also coloring the seeds of the open experiment. This was done to avoid that any harvest difference could be attributed to the colorant itself.

where α is the expected output with the local seed, $E(Y_{0i})$,and η_i is the random part of Y_{0i}, $\eta_i = Y_{0i} - E(Y_{0i})$. Expected outputs with improved and local seeds are respectively given by:

$$\begin{aligned} E(Y_i|\tau_i = 1) &= \alpha + \rho + E(\eta_i|\tau_i = 1) \\ E(Y_i|\tau_i = 0) &= \alpha + E(\eta_i|\tau_i = 0) \end{aligned} \tag{4}$$

As the allocation to the improved seeds and local seeds groups has been artificially randomized, we have:

$$E(Y_{i0}|\tau_i = 1) - E(Y_{i0}|\tau_i = 0) = 0 \tag{5}$$

And the difference in sample means yields:

$$E(Y_i|\tau_i = 1) - E(Y_i|\tau_i = 0) = \rho \tag{6}$$

RCTs offer hence a direct way of estimating the impact of improved seeds on output, bypassing the traditional pitfall of self-selection bias, a threat to the internal validity of impact studies of improved seeds adoption based on observational data[27].

Rational farmers should adjust farm management if the improved seeds change the marginal product of the inputs at their disposal (e.g. Bulte et al. 2014). For example, farmers should either work harder, re-allocate household labour from other activities towards the targeted plot, hire additional labour or apply more fertiliser. There are hence two channels by which improved seeds increases output: the

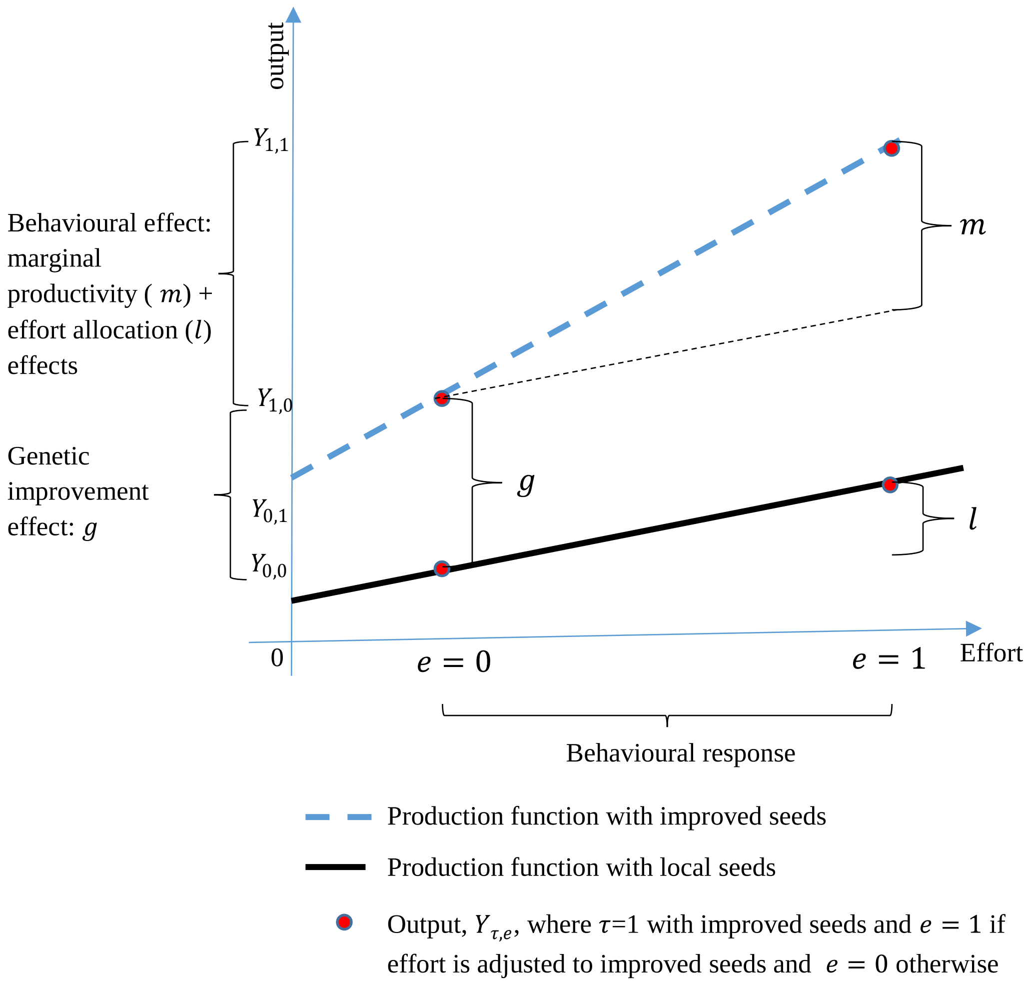
[27] If we were to compare outputs between adopters and non-adopters with observational data, i.e. once farmers have decided to adopt or not, a simpler comparison of means would yield the following:

$$\begin{aligned} E(Y_i|\tau_i = 1) - E(Y_i|\tau_i = 0) &= \rho + E(\eta_i|\tau_i = 1) - E(\eta_i|\tau_i = 0) \\ &= \rho + \underbrace{E(Y_{i0}|\tau_i = 1) - E(Y_{i0}|\tau_i = 0)}_{Selection\ bias} \end{aligned} \tag{i}$$

The selection bias is caused by a systematic difference in the conditional output under non-adoption of the adopters and the non-adopters. In the case of improved seeds, we can typically expect adopters to be more productive than non-adopter even before adoption. In the regression based on equation (3), the dummy variable τ_i is hence correlated with η_i so that an Ordinary Least Squares (OLS) estimate of ρ would be biased upward. The selection bias has notably been dealt with endogenous switching regressions and propensity score matching methods. In the first approach, the decision to adopt and the impact of adoption on output are analysed jointly. The identification of the causal effect of adoption on the yield rests on the assumption that at least one variable used to explain adoption is not correlated with yield (exclusion restriction). In the propensity score matching methods, a set of observable variables is used to build a comparable set of pairs of adopters and non-adopters. The assumption is that conditional on these variables, the decision to adopt is random. However, some characteristics determining adoption might be unobservable to the analyst, implying that the conditional independence assumption is violated. The advantage of randomly allocating farmers to either the improved or local seeds group is that it solves the selection bias issue. As farmers are randomly allocated to groups receiving either the improved or local seeds, the expected output conditional on non-adoption is equal for both groups.

higher genetic potential of the improved seeds, the behavioural response. We illustrate each effect in Figure 3.1.

Figure 3.1: The effects of technology adoption



Effort, e, is plotted on the horizontal axis of Figure 3.1 and output, $Y_{\tau e}$, on the vertical axis, where $\tau = 1$ with improved seeds and zero with local seeds and $e = 1$ if effort is adjusted to the improved seeds and $e = 0$ otherwise. The improved seeds lead both to a vertical shift of the production function and an increase in its slope.

We define the *genetic improvement effect*, g, as the increase in output at the default level of effort, $Y_{10} - Y_{00}$. It is typically the effect of improved seeds if farmers face input constraints preventing them from adjusting their effort from $e = 0$ to $e = 1$.

We define the *effort allocation effect, l*, as the increase in output caused by the increase in effort alone, i.e. in the increase in output that could have been obtained with local seeds given an increase in effort from $e = 0$ to $e = 1$.[28] The effort allocation effect is the difference between Y_{01} and Y_{00}.

Lastly, we define the *marginal productivity effect, m,* as the increase in output at the optimal level of effort ($e = 1$) caused by the increase in the productivity of effort. It is given by $(Y_{11} - Y_{10}) - (Y_{01} - Y_{00})$.

The potential output model of equation (1) can be extended to:

$$Y_i = \begin{cases} Y_{11i} & \text{if } \tau_i = 1 \text{ and } e_i = 1 \\ Y_{10i} & \text{if } \tau_i = 1 \text{ and } e_i = 0 \\ Y_{01i} & \text{if } \tau_i = 0 \text{ and } e_i = 1 \\ Y_{00i} & \text{if } \tau_i = 0 \text{ and } e_i = 0 \end{cases} \tag{7}$$

where $e_i = 1$ if the farmer increase effort and $e_i = 0$ otherwise, $\tau_i = 1$ if the farmers received improved seeds and $\tau_i = 0$ otherwise. The four counterfactuals correspond to the following possibilities: the farmer received improved seeds and increases effort, Y_{11i}, the farmer received improved seeds and does not increase effort, Y_{10i}, the farmers received local seeds and increases effort, Y_{01i}, the farmers received local seeds and does not increase effort, Y_{00i}.

This can be expressed as:

$$\begin{aligned} Y_i = Y_{00i} + (Y_{10i} - Y_{i00})\tau_i + (Y_{01i} - Y_{i00})e_i \\ + \big(Y_{11i} - Y_{i10} - (Y_{01i} - Y_{i00})\big)\tau_i e_i \end{aligned} \tag{8}$$

Assuming a constant effect model and letting $Y_{10i} - Y_{00i} = g$, $(Y_{01i} - Y_{i00}) = l$, $\big(Y_{11i} - Y_{i10} - (Y_{01i} - Y_{i00})\big) = m$, we can express equation (9) as:

$$Y_i = \alpha + g\tau_i + le_i + e_i\tau_i m + \eta_i \tag{9}$$

where $\alpha = Y_{00i}$, g is the genetic improvement effect, l is the effort allocation effect, m is the marginal productivity effect and $\eta_i = Y_{00i} - E(Y_{00i})$.

The total effect of improved seeds, ρ in equation (6) can hence be decomposed as:

$$E(Y_i|\tau_i = 1, e = 1) - E(Y_i|\tau_i = 0, e = 0) = g + l + m \tag{10}$$

[28] It is therefore called the pseudo-placebo effect in Bulte et al. (2014) following the terminology of Chassang et al. (2012).

If a large part of the total effect is driven by the behavioural effect ($l + m$), the results of the RCT might not hold in another region or in another population sub-group because of differences in input constraints and differences in the perceived benefits of the new technology for instance, i.e. the RCT might have a low external validity.

Furthermore, if the interest lies in the increase in output caused by the higher productivity of the improved seeds, i.e. the net effect of the improved seeds, the effort allocation effect, l, should be taken out of the equation (10) as it could occur as well with local seeds. The net effect can hence be expressed as:

$$net\ effect = g + m \tag{11}$$

In Figure 3.1, the net effect is given by $Y_{11} - Y_{01}$.

3.3 Identification strategy

There are several possible routes to isolate each effect. An option is to run a regression where we control for effort, e, and interact it with the improved seeds dummy, expressing equation (9) as:

$$E(Y|e,\tau) = \alpha + g\tau + le + m\tau e \tag{12}$$

where e is now a proxy variable for effort and $\tau = 1$ if the farmers received improved seeds and zero otherwise. Even in the relatively simple setting of the production function of smallholders in developing countries, controlling for effort might be hard as inputs vary in quantity and quality; effort might be adjusted across several dimensions (Giller et al. 2011). A large set of proxy variables should be used and interacting them with m would greatly increase the number of parameters to estimate, reducing hence the power of the test we would like to carry on each parameter.

Our solution is to opt for an experimental design similar to Chassang et al. (2012) whereby both the improved seeds and the probability of receiving them are randomized.

We summarize the design of the experiment in Table 3.1**.** Participants were allocated to either of four groups.

Table 3.1: Design of the experiment

	Improved seeds	**Local seeds**
Open RCT	G1	G2
Double-blind RCT	G3	G4

The two first groups were allocated to a traditional open RCT experiment: half received improved seeds (G1) and the other half got local seeds (G2). Both groups were told about the type of seeds they got. In other terms, farmers in G1 were told they had a probability of 100% of receiving the improved seeds and farmers in G2 were told they had a probability of 0% of getting the improved seeds.

The third and fourth groups were allocated to a double-blind RCT: half received improved seeds (G3) and the other half got local seeds (G4). G3 and G4 were told they had one chance out of two of getting the improved seeds, i.e. a probability of 50%.

Our identification strategy rests on two assumptions. First, effort is weakly increasing in the probability of receiving improved seeds:

$$e_i(0) \leq e_i(0.5) \leq e_i(1) \qquad \text{A1}$$

where $e_i(p)$ is the effort of a farmer on the experimental plot and is both a function of his type i and the probability of having received improved seeds p where p has been communicated to him by the extension agent in charge of distributing the seeds. We assume hence that effort is weakly increasing in p.

Second, there are no Hawthorne or John Henry effects[29]: the behavioural response to the probability of receiving improved seeds is not caused by participants' awareness of taking part in an experiment. It is rather a response to the expected increase in marginal productivity of effort caused by the improved seeds and would also take place independent of an experimental setting.

Let us rewrite the potential output model of equation (7) as:

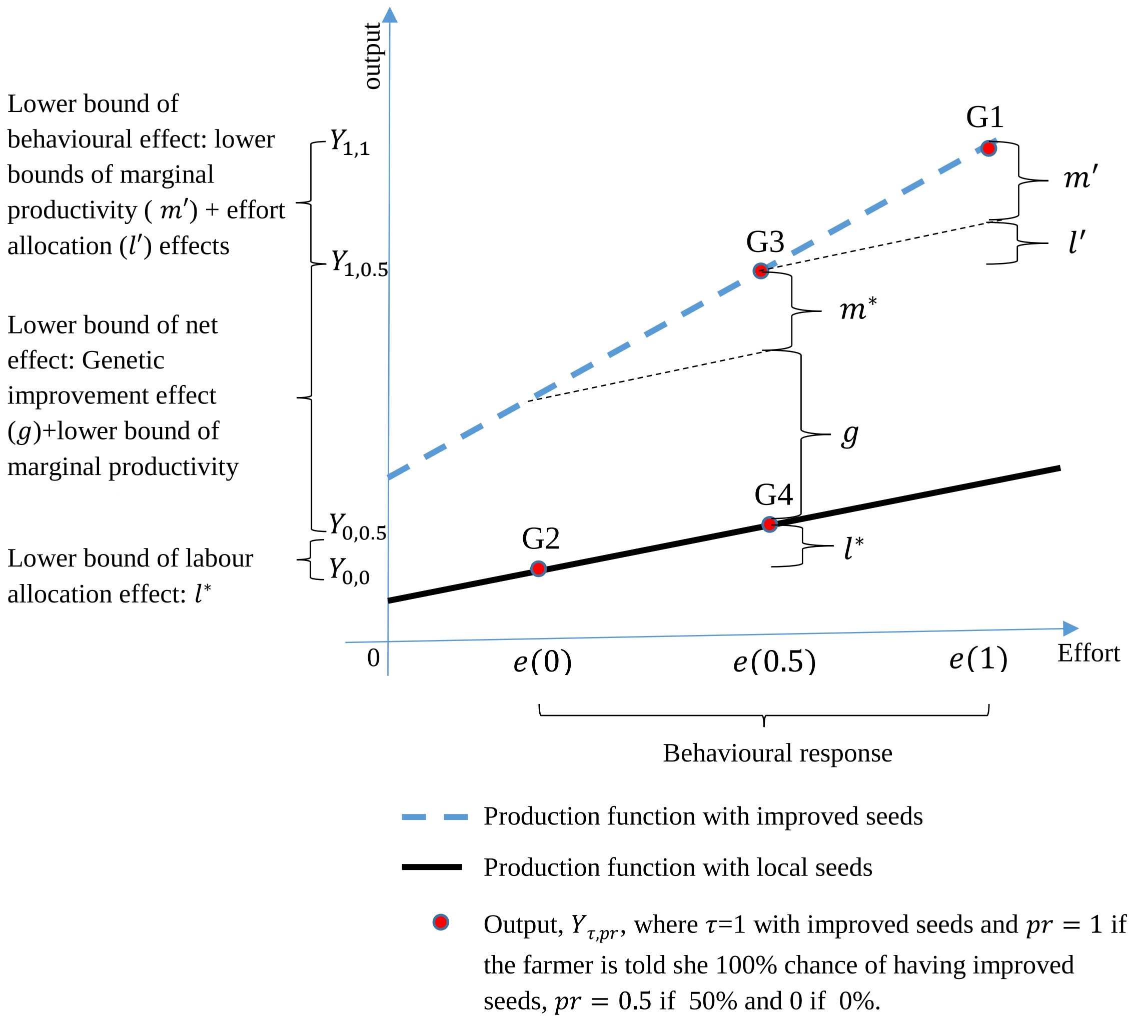
$$Y_{i,\tau,p} = \begin{cases} \mu_i\big(1, e_i(1)\big) + \varepsilon_{i,1,1} & \text{if } \tau_i = 1 \text{ and } p_i = 1 & \text{Group 1} \\ \mu_i\big(0, e_i(0)\big) + \varepsilon_{i,0,0} & \text{if } \tau_i = 0 \text{ and } p_i = 0 & \text{Group 2} \\ \mu_i\big(1, e_i(0.5)\big) + \varepsilon_{i,1,0.5} & \text{if } \tau_i = 1 \text{ and } p_i = 0.5 & \text{Group 3} \\ \mu_i\big(0, e_i(0.5)\big) + \varepsilon_{i,0,0.5} & \text{if } \tau_i = 0 \text{ and } p_i = 0.5 & \text{Group 4} \end{cases} \qquad (13)$$

where $\mu_i\big(\tau, e_i(p)\big)$ is the expected output for farmer of type i under treatment status τ and effort $e_i(p)$. The error $\varepsilon_{\tau p i}$ represents production shocks such as water deficit or excessive rainfall, pest invasion or plant disease. It has expectation $E\big(\varepsilon_{\tau p i}|i\big) = 0$. Each type i summarizes all observed and

[29] The Hawthorne effect describes a situation in which participants allocated to the treatment group increase their effort because they know they are assigned to the treatment group. The John Henry effect describes the same situation but for those allocated to the control group.

unobserved factors affecting output. In our case, these could comprise farm management skills, access to credit and inputs, and farmers' readiness to adopt new farming practices.

Figure 3.2: Effects estimated in the experiment



The comparison in means in the open RCT (groups 1 and 2), provides the average total effect of the improved seeds:

$$E(Y_i|\tau_i = 1, p_i = 1) - E(Y_i|\tau_i = 0, p_i = 0) = \hat{g} + \hat{l} + \hat{m} \qquad (14)$$

It is illustrated in Figure 3.2 as the difference between G1 and G2 and was denoted ρ in equation (6).

By comparing output in the double-blind RCT (G3-G4), we get a lower bound of the net effect of improved seeds:

$$E(Y_i|\tau_i = 1, p_i = 0.5) - E(Y_i|\tau_i = 0, p_i = 0.5) = \hat{g} + \hat{m}^* \tag{15}$$

where $\hat{m}^* \leq \hat{m}$ because of assumption (A1).

The lower bound of the net effect of adoption is illustrated in Figure 3.2 as the difference between G3 and G4.

By comparing the average treatment effect in the open and double-blind RCTs, we get upper and lower bounds of the net effect of the improved seeds adoption:

$$\hat{g} + \hat{m}^* \leq \hat{g} + \hat{m} \leq \hat{g} + \hat{m} + \hat{l} \tag{16}$$

Let us now compare participants with the same seeds, but with different probabilities of receiving improved seeds. Comparing group 4 and group 2, i.e. farmers with local seeds in the double-blind and open RCTs, we obtain a lower bound of the effort allocation effect:

$$E(Y_i|\tau_i = 0, p_i = 0.5) - E(Y_i|\tau_i = 0, p_i = 0) = \hat{l}^* \tag{17}$$

where $\hat{l}^* \leq \hat{l}$ because of assumption (A1). The lower bound of the effort allocation effect, $\hat{l}^*$, is illustrated in Figure 3.2 as the difference between G4 and G2.

Lastly, by comparing group 1 with group 3, i.e. farmers with improved seeds in the open and double-blind RCTs, we get a lower bound of the behavioural effect:

$$E(Y_i|\tau_i = 1, p_i = 1) - E(Y_i|\tau_i = 1, p_i = 0.5) = \hat{l}' + \hat{m}' \tag{18}$$

where $\hat{l}' \leq \hat{l}$ and $m' \leq \hat{m}$ because of assumption (A1). The lower bound of the behavioural effect is illustrated in Figure 3.2 as the difference between G1 and G3.

By comparing output between groups, we can hence get an estimate of the total impact of improved seeds, $\hat{g} + \hat{l} + \hat{m}$, a lower bound estimate of the net impact, $\hat{g} + \hat{m}^*$, a lower bound estimate of the effort allocation effect, $\hat{l}^*$, and lower bound of the behavioural effect, $\hat{l}' + \hat{m}'$. Table 3.2 summarises the effects estimated in the experiment.

Table 3.2: Summary of the effects estimated in the experiment

Groups	Effects	Expression
G1-G2	Total effect	$\hat{g} + \hat{l} + \hat{m} = E[\mu_i(1, e_i(1)) - \mu_i(0, e_i(0))]$
G3-G4	Lower bound of net effect	$\hat{g} + \hat{m}^* = E[\mu_i(1, e_i(0.5)) - \mu_i(0, e_i(0.5))] \leq \hat{g} + \hat{m}$

G4-G3	Lower bound of labour allocation effect	$\hat{l}^* = E\left[\mu_i\left(0, e_i(0.5)\right) - \mu_i\left(0, e_i(0)\right)\right] \leq \hat{l}$
G1-G3	Lower bound of behavioural effect	$\hat{p}^* + \hat{b}^* = E\left[\mu_i\left(1, e_i(1)\right) - \mu_i\left(1, e_i(0.5)\right)\right] \leq \hat{p} + \hat{b}$

The model above rests on the assumptions that effort is weakly increasing in the probability of receiving improved seeds and that there are no Hawthorne or John Henry effects. While we can test the first assumption by investigating if effort does vary between groups, we cannot test the second assumption. In order to identify separately the Hawthorne effect and the John Henry effect from the effects described above, we would need a third treatment arm where famers would have been allocated randomly to more or less intensive follow-up and scrutiny from the research team (e.g. McCarney et al. 2007). Lastly, in order to obtain direct estimates rather than lower bound estimates of the net, effort allocation and behavioural effects, a different experiment could have been ran. Farmers with improved seeds in the double-blind RCT should have been told they had local seeds and vice-versa. However, it would have raised important ethical questions. This is the reason why we preferred the experimental design discussed above.

3.4 Data

Maize, along with rice and wheat, is one of the three major crops comprising 70% of world food production, and plays a major role in the diet of Sub-Saharan populations. There has been a great deal of research effort directed toward improving maize yields since the start of the green revolution in the 1960s. We can therefore expect to see large increase in yields with modern maize varieties. The improved seeds tested in the current study, the Situka-M1, was released in 2001 by Selian Agricultural Research Institute(SARI). It has a yield potential of 3-5 ton/ha and its optimal production altitude ranges from 1000 to 1500m above sea level. In Tanzania, it can grow in the Eastern and Northern regions where our study areas are located. The variety is tolerant to drought, maize streak and grey leaf spot diseases, and resistant to *Diplodia* fungus, *Fusarium* leaf bright and *Puccinia sorghi*. Although its yields are often advertised as 4 to 6 ton/ha by the government (Ministry of Agriculture, Food Security and Cooperatives. 2009, cit. in Tumbo et al. 2012) or grain dealers (e.g. Suba Agro-Trading & Engineering Co. Ltd[30]), CIMMYT found considerably lower yields, from 2.4 ton /ha in a mid-altitude dry environment to 4 ton/ha in a mid-altitude humid hot environment (Magorokosho et al. 2009). Naturally, yield does not depend only on the type of seed: input choices, soil quality and weather conditions are also key determinants. A given type of seeds is therefore expected to perform differently according to the farmer's skills, his access to inputs and the weather.

[30] Website accessed on the 6th of December 2013: http://subaagro.com/index_files/OPV.htm.

Agriculture is the main employment sector in Tanzania and represented 80% of the labour force in 2003 (Eskola 2005). Maize is the major staple crop produced and consumed in Tanzania (Amare et al. 2012; Shiferaw et al. 2014). It accounts for 45% of the cultivated land and 75% of cereal production (Kassie et al. 2014). Virtually all the maize is produced by smallholders (98%, Minot 2010), with 0.8 ha per farm on average dedicated to maize. Although the land under cultivation increased by 54% between 2000 and 2015 (Kassie et al. 2014), only 18% of land planted with maize was planted with improved maize seeds in 2006 and similar figures still apply in the study area (Kassie et al. 2014; Smale et al. 2011). Yields are low compared to Europe: 1.5 ton per ha against more than 9 ton per ha in the European Union (FADN 2013). The average per capita annual maize consumption is 73kg, contributing on average to 33% of calorific needs (Minot 2010). Much of the output is consumed on the farm (Alene et al. 2009), so the benefit of improved seeds adoption might be assumed to imply higher household food consumption.

The data are based on two sets of RCTs ran in parallel in the 2013 main growing season. The farmers were spread in three districts of Tanzania (Karatu, Mvomero and Kilosa), covering the main agri-ecological zones of Tanzania. Karatu, in the northern part of Tanzania, is located next to the natural Ngorongoro conservation area and to the tarmac road which brings numerous visitors each year to the Serengeti national park. Despite the proximity of this tourist attraction, farmers in the surrounding villages do not benefit much from this flow of travellers as most do not stop in Karatu. The 399 farmers who took part to the experiment in Karatu district live in three villages that are within a maximum of 20 km of each other. Despite their relative proximity, each one belongs to a distinct agro-ecological zone: Changarawe is located at an altitude of 1350m-1450m with a dry climate; Kilimatembo and Rothia benefit from wetter conditions and are located at an altitude of 1500m-1600m and 1600m-1700m respectively. The 290 farmers who took part to the study in the East are spread over two districts (Kilosa and Mvomero) and 12 villages. By contrast to the Karatu area, there are no tourist activities and these villages are far more remote from one another - the maximum distance between each one being close to 140 km. They are located at a lower altitude, between 500m and 1075m and are diverse in terms of humidity. Most are distant from any tarmac road and the closest village to the regional centre, Morogoro, is still 25 km away from it.

Kassie et al. found (2014) that one fifth of the farmers adopted improved maize seeds in the study area (20% in Karatu, 25% in Kilosa and 17% Mvomero) while Amare et al. (2012) report an adoption rate of 50% in Karatu. Maize accounts on average for 70% of crop production and constitute 80% of domestic food production consumption in the study area (Kassie et al. 2014). Kassie et al. (2014) found yields of 1.2 t per ha for adopters of improved maize varieties compared to 0.5 t. per ha for local varieties.

At the beginning of December 2012, farmers were told by extensions agents, that they would have the opportunity to take part to a study on maize yield. Seeds were distributed in January in units of circa 2 kg (2 tins). Farmers were allocated randomly to four groups. Group 1 got Situka M1 seeds and group 2, local seeds. Farmers in groups 1 and 2 were told the type of seeds they were using, as in a classical open RCT. Farmers in groups 3 and 4 got Situka M1 and local seeds respectively but were not told to which category they were assigned nor were the extensions agents in charge of distributing the seeds. It is hence a double-blind experiment. The Situka M1 is treated with a pink fungicide powder while local seeds are not. In order to make them indistinguishable, we applied a reddish food colorant on both types of seeds. Farmers in the double-blind groups were told by extension agents that they had one chance out of two of getting a bag with improved seeds.

Furthermore, farmers in all four groups were told that a reddish colorant had been applied on all types of seeds and that they could plant the seeds and manage the plots as they wanted. The size of the plot, the soil quality (farmers have multiple small plots of various sizes and qualities), the number of seeds per hole, the spacing between rows as well as the number of weeding and threshing are all important production choices.

Co-authors returned six times for two-week stays to monitor the progress of the experiment and collect data at different stages of the growing season, the final survey being conducted in July and August in the eastern and northern districts respectively. Most farmers in the double-blind groups found out the true type of seeds at maturation because the Situka M1 is an early maturing maize breed. However, as most production choices were already made at this stage of the growing cycle, this is of minor concern for the results.

The attrition rate was limited (9%). Among the 625 farmers having answered the end-line survey, we identified 10% of non-compliers, i.e. farmers in the double blind groups who were told which type of seeds they had by the extension agents, farmers who were in the open RCT but were not told the type of seeds they got, and farmers who got the wrong type of seeds according to the randomization scheme. The final sample therefore includes 560 farmers, of which 348 are located in the North and 212 in the East.

We tested if randomization worked by comparing a set of 20 pre-determined socio-economic variables such as farm size, gender of the household head, education, and others. The only statistical difference we found was the average age of group 3 farmers which was slightly lower than group 4 farmers. The difference is nevertheless pragmatically negligible (45 years old v. 48 years old) and likely to be driven only by a few outliers (see Table 3.9 in Appendix 3.8). We hence conclude that randomization worked. We further tested if non-compliant farmers differ according to this set of 20 variables and failed to reject the null hypothesis that they do not differ from the rest of the sample. Nevertheless, we choose to remove them from the sample. Note that the main results of the present study hold true when tested on the sample including non-compliers. The effect estimated is hence the Local Average Treatment Effect, i.e. the effect on the compliers only. However, given the good balance in terms of socio-economic characteristics, the results are likely indicative of the Average Treatment Effect (ATE), a term which we will use below when commenting on the estimated effects.

Table 3.3 presents a set of summary statistics. The sample average harvest on the experimental plot is 95 kg, the average yield is 0.82 ton per ha (ton/ha), and the average plot size is 0.11 ha. The standard deviation of yield is very large at harvest (ton/ha) and the minimum yield is 0 kg in case of crop failure and the maximum is 5.53 ton/ha. The top values likely represent enumeration errors: farmers might have provided the total harvest quantity on their farm rather than on the experimental plot. Crop failures were mostly caused by termite infestation in the East and excessive rainfall in the North (some steep plots were washed away by heavy rains).

The median farm size is 1.2 ha, with a maximum of 8.5 ha, the 75th percentile at 2 ha and the 95% percentile at 4 ha. Farmers are hence mostly small-scale subsistence farmers with only a small fraction of the sample likely to have important and regular surplus for sale. The average area planted with maize measures around 0.5 ha, which means that we provided the seed for close to a quarter of the farm's total maize production. The experimental plot was hence an important part of the production process. As farmers were told that we would not take any of the harvest, it is likely that they have been managing it with no less care than their usual plots.

Only one third of the experimental plots have good quality soil, less than half are flat, while close to 40% of them suffer from soil erosion according to plot measurements done by extension agents. Hence, plots tend to be of lesser quality than the ideal plots on which yield estimates are based when marketing a new maize breed. Furthermore, field visits of more than 50 plots showed great differences in management, in terms of planting decisions and weeding, for instance.

Table 3.3: Summary statistics

	Mean	St. dev.	Median
Harvest (kg)	95.23	105.49	60.00
Yield (ton/ha)	0.82	0.85	0.59
Size of the plot (ha)	0.11	0.05	0.10
Good soil	0.35	0.48	0.00
Flat	0.46	0.44	0.46
Erosion	0.39	0.43	0.39
Labour (man days)	8.79	6.64	7.40

One or no weeding	0.18	0.38	0.00
Fertiliser (mostly manure)	0.22	0.42	0.00
Pesticide	0.11	0.32	0.00
Plot improvement work (soil bund, terrace etc.)	0.29	0.46	0.00
Intercropping	0.14	0.31	0.00
Standardized precipitation Index (peak rainfall month, ARC 2 dataset)	0.14	0.39	0.23
Extreme precipitation (self-reported)	0.48	0.50	0.00
Crop damage due to pest, disease or fungi	0.26	0.44	0.00
Female headed household	0.12	0.33	0.00
Age of household head	46.22	12.93	46.22
Education: one household member up to form IV	0.18	0.39	0.00
Household size	5.14	2.21	5.00
Dependency ratio (no of dependent per adult)	0.55	0.68	0.33
Risk lover (risk experiment)[31]	0.41	0.49	0.00
Land owned (ha)	1.22	1.18	0.91
Land farmed (ha)	1.49	1.22	1.21
Motorbike	0.15	0.36	0.00
Rich (self-reported)	0.26	0.44	0.00
Oxen	0.26	0.44	0.00
Active role in the community	0.27	0.44	0.00
Member of a self-help group (Sacco, vicoba, funeral society)	0.38	0.49	0.00
Member of a social association (e.g. reglious, youth, women)	0.74	0.44	1.00
Social network for help in cash, in kind or on the farm (number of people)	2.67	3.51	2.00
Social network for agricultural related information (number of people)	1.77	3.57	1.00

In terms of labour, households have allocated on average 9 man-days to the experimental plot although there are some outliers, probably driven by the same kind of enumeration errors as mentioned above. One quarter of households own at least one ox, an important input for land preparation, most of them in the North.

The average household size is 5 people with an average dependency ratio of 50% (one adult for two children). Female-headed households represent a substantial minority (12%). In terms of social capital, close to 30% of the households count at least one member in a village community organisation, 40% are member of a credit union or other self-help group, and 70% are a member of a social group such as youth, women or religious associations. We have two proxies for social network: the number of

[31] Farmers were presented with a set of six lotteries ranked according to their riskiness (variance) and expected gain. Once the farmer had chosen his preferred lottery, a coin was flipped and he received the corresponding pay-off. The maximum amount was 4000 Tsh (circa 2 USD). Farmers who took one of the 2 riskiest lotteries were classified as risk lover seeking (40% of the sample). See appendix B for the protocol.

people the farmer can ask for help in cash, in kind or on the farm and the number of people he can ask agriculture-related questions. The median size of the help and information networks is 2 and 1 people, respectively.

Table 3.4 show large differences in key inputs between regions. In the East, farmers worked 50% more on their plot than farmers in the North, planted the crop on larger plots, and a greater share of the plots were flat with good soils and no erosion. Furthermore, close to 30% of the farmers located in the East chose one of their best plots for the experiment while the share drops to 10% in the North.

Fertilisers were used more often in the North (mostly manure), which is likely driven by the much higher share of households holding cattle in the North (40% against 2% in the East). The lower use of manure in the East is hence not the result of a choice but of input constraints. Pesticides were used more often in the North, suggesting a greater threat from pest damage although actual pest damage does not differ between regions but plot visits suggested more termite attacks in the East. Experimental plots were more likely to have benefited from improvement work in the North, although most of the improvement work is aimed at limiting soil erosion (soil and rock bunds, terracing etc.), which is needed less in the East given the flatness of the terrain. Mulching was for instance carried in similar proportion in both regions. With the exception of manure, it appears hence that conditions and farm management has been more favourable in the East than in the North.

Table 3.4: Differences in key inputs between regions, comparison in mean

	East	North	East-North
Labour (man day)	7.42	4.88	2.54^{***}
	(3.64)	(3.21)	(0.30)
Plot size (ha)	0.12	0.10	0.01^{***}
	(0.04)	(0.05)	(0.00)
Pesticide use (d.)	0.02	0.17	-0.15^{***}
	(0.15)	(0.38)	(0.03)
Fertiliser (, d.)	0.05	0.33	-0.29^{***}
	(0.21)	(0.47)	(0.03)
Good soil (d.)	0.44	0.30	0.15^{***}
	(0.50)	(0.46)	(0.04)
More fertile plot (d.)	0.28	0.09	0.19^{***}
	(0.44)	(0.29)	(0.31)
Flat plot (d.)	0.67	0.31	0.37^{***}
	(0.47)	(0.46)	(0.04)
Erosion (d.)	0.29	0.45	-0.15^{***}
	(0.46)	(0.50)	(0.05)
Plot improvement (d.)	0.21	0.34	-0.14^{***}
	(0.41)	(0.48)	(0.04)
One or no weeding (d.)	0.19	0.17	0.01
	(0.39)	(0.38)	(0.03)

p-values of the t-test: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Labour is the pre-harvest labour (land preparation, planting, weeding). d. stands for dummy variables.

GPS coordinates were taken for each village and were used to match the survey data with the rainfall data from the African Rainfall Climatology Version 2 dataset. Weather conditions have been on average slightly wetter than normal according to the standardized precipitation index (SPI, Guttman 1999). The average SPI value for March precipitation, the month during which most of the rain falls, is indeed 0.27, so lies between the bounds of nearly normal conditions (-0.5, 0.5). Some villages in the East, Manza and Vitonga, were exposed to relatively dry conditions with a return period of approximately 10 years (SPI values of -1.27) while Kilimatembo (North) enjoyed quite wet conditions (SPI 0.9, return period of approximately 5 years). According to informal discussions with farmers and extension agents in the North, the timing of rain was not optimal this year: rain started very late, then fell so heavily during one month that some plots were washed away, and then stopped abruptly, leaving maize plants short of the optimal rainfall required for their growing phase.

3.5 Results

We start by comparing effort level across treatment groups in order to investigate if assumption A1, i.e. effort is weakly increasing in the probability of receiving improved seeds, does hold. We then present the average yields[32] across groups and regions and proceed to the estimation of the total effects, the lower bounds of the net effects, the lower bounds of the labour allocation effects and the lower bounds behavioural effects. We left out of the analysis the top 5% of the yield distribution in order to limit the effect of outliers.

3.5.1 Behavioural response to the probability of receiving improved seeds

Table 3.5 shows the average behavioural response to the change in the probability of receiving improved seeds. Under assumption A1, effort should be weakly increasing in the probability of obtaining improved seeds. Therefore, average effort among farmers with a probability of 100% of getting improved seeds (group 1) should be higher or equal to the average effort of those with a 50% probability of getting improved seeds (groups 3 and 4), which , in turn, should be higher or equal to the average effort of the farmers with a zero probability of getting improved seeds (group 2).

We used 9 variables in order to capture effort: pre-harvest labour (man days)[33], size of the experimental plot (ha), pesticide use (dummy), fertiliser use (dummy), good soil (dummy), flat plot

[32] We chose to present the results in terms of yields rather than harvests because it is a more common metric. Results do hold as well when using harvests instead of yields.

[33] We want here to measure the effort allocation causing a harvest and not effort allocation caused by a harvest. We therefore focus on the pre-harvest labour.

(dummy), erosion (dummy), plot improvement (dummy), one or no weeding (dummy). Results are shown in Table 3.5.

In the North, we cannot reject the null hypothesis of no difference in effort between groups: there appears to be no behavioural response to the probability of receiving improved seeds. According to the model presented in section 3.3, it implies that we should not find any significant difference between the open and the double-blind RCT and, as a results, the lower bound of the behavioural effects should not be significantly different from zero.

Table 3.5: Behavioural response to the probability of receiving improved seeds

	North			East		
Difference in probability of receiving improved seeds :	100%-0% (G1-G2)	100%-50% (G1-DB)	50%-0% (DB-G2)	100%-0% (G1-G2)	100%-50% (G1-DB)	50%-0% (DB-G2)
Labour (man day)	-0.05	0.51	-0.56	1.18*	0.96	0.22
	(0.76)	(0.59)	(0.62)	(0.63)	(0.72)	(0.65)
Plot size (ha)	0.01	0.01	0.00	0.02**	0.02**	0.00
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Pesticide^D	-0.07	-0.01	-0.06	0.01	0.02	-0.02
	(0.06)	(0.05)	(0.05)	(0.03)	(0.03)	(0.02)
Fertiliser^D	-0.02	-0.05	0.03	0.03	0.03	-0.01
	(0.07)	(0.06)	(0.06)	(0.04)	(0.04)	(0.03)
Good soil^D	0.03	-0.02	0.05	0.03	0.03	0.01
	(0.07)	(0.06)	(0.06)	(0.09)	(0.09)	(0.08)
Flat plot^D	-0.12	-0.05	-0.07	-0.04	-0.07	0.02
	(0.08)	(0.07)	(0.07)	(0.09)	(0.09)	(0.08)
Erosion^D	0.07	0.01	0.06	0.10	0.07	0.03
	(0.08)	(0.08)	(0.07)	(0.09)	(0.09)	(0.08)
Plot improvement^D	0.05	0.02	0.03	-0.09	-0.05	-0.04
	(0.07)	(0.06)	(0.06)	(0.07)	(0.07)	(0.07)
Few weeding^D	-0.09	-0.06	-0.03	-0.02	-0.03	0.00
	(0.06)	(0.05)	(0.05)	(0.07)	(0.07)	(0.06)

G1*: Improved seeds in the open RCT (probability of improved seeds=100%),* ***G2****: Local seeds in the open RCT (probability of improved seeds=0%),* ***DB****: double-blind groups (groups 3 and 4) with a probability of 50% of receiving the improved seeds. p-values of the t-test:* * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. *Labour is the pre-harvest labour (land preparation, planting, weeding).* ^{D} *stands for dummy variables.*

In the East, by contrast, we do find significant differences in the pre-harvest labour and the size of the experimental plot. As expected, farmers with a 100% probability of receiving improved seeds (group 1) allocated on average more labour to the plot and chose a larger plot than those with a 0% probability of receiving improved seeds (group 2). Similarly, farmers with a 50% probability of getting the improved seeds (group 3 and 4) chose a larger plot on average than those with a 0% probability of getting improved seeds (group 2). Although the difference in terms of labour allocation is not significant between the groups with a 100% and 50% chance of getting improved seeds (group 1 and the double-blind groups) and between those with a 50% and 0% (double-blind groups and group 2), it has the expected sign: the higher the probability of receiving improved seeds, the higher is the average labour allocation. We can therefore expect to find a difference in the results of the open and double-blind RCTs and a significant behavioural effect.

Assumption (A1) appears hence to hold across the sample: effort is weakly increasing in the probability of receiving improved seeds. We expect that the difference in the behavioural response between regions should be reflected in the results of the experiments: the larger behavioural response in the East should imply a larger behavioural effect.

3.5.2 Average treatment effects

Table 3.6 shows the average yields across treatment groups and regions. Farmers in the East have on average higher yields than those in the North. This is consistent with the observation that plots were of better quality in the East (larger, better soil, lower erosion, flatter etc.) and the higher labour invested on the plots (Table 3.4). However, yields are very low in both cases: 0.59 ton/ha in the North and 1.16 in the East, less than half those found in Magorokosho et al. (2009). In both regions, improved seeds have a large impact on average yield: they are close to 60% higher than with local seeds. In terms of the second treatment arm of the experiment, i.e. the probability of receiving improved seeds, we observe that the average yields with improved seeds in the open category (G1) tend to be higher than in the double-blind category (G3). By contrast, farmers having received local seeds in the open RCT (G2) and the double-blind RCT (G4) have similar yield levels. We test below if these differences are statistically significant (Table 3.7).

Table 3.6: Average harvest and yields

	Yield (ton/ha)						
	All	Local G2 & G4	Improved G1 & G3	G1	G2	G3	G4
Both regions	0.82	0.64	1.01	1.12	0.69	0.87	0.57
	(0.85)	(0.58)	(1.03)	(1.16)	(0.61)	(0.81)	(0.53)
North	0.59	0.46	0.73	0.76	0.46	0.69	0.46
	(0.51)	(0.36)	(0.62)	(0.66)	(0.35)	(0.57)	(0.37)
East	1.16	0.91	1.43	1.65	0.96	1.13	0.82
	(1.10)	(0.72)	(1.34)	(1.51)	(0.72)	(1.03)	(0.71)

***G1**: Improved seeds in the open RCT, **G2**: Local seeds in the open RCT, **G3**: Improved seeds in the double-blind RCT, **G4:** Local seeds in the double-blind RCT. Standard deviations in parentheses.*

Following the model presented in section 3.3, the results of the open RCT, i.e. the comparison in mean values between groups 1 and 2, provide the average total effect of improved seed: $\hat{g} + \hat{l} + \hat{m}$. As expected, the total effect is large. In the North, the average yields with the local and improved seeds are, respectively, 0.46 ton/ha and 0.76/ha, i.e. an increase of 0.31 ton/ha (or 65%) significant at the 99% confidence level. In the East, average yield with local and improved seeds are, respectively, 0.96 ton per ha and 1.65 ton/ha, i.e. an increase in yields of 0.69 ton/ha (or 72%). Lastly, at the aggregate level, yields increase by 0.44 ton/ha (62%). The total effect is hence very large both in the North and the East.

Table 3.7: Comparison in means

Treatment Effects on Yield (ton/ha)		East	North	Both regions
Total effect: G1- G2	g+l+m	0.69***	0.31***	0.44***
		(0.21)	(0.08)	(0.11)
LB of Net effect: G3-G2	g+m*	0.31	0.23***	0.29***
		(0.2)	(0.08)	(0.09)
LB of Behavioural effect: G1-G3	l'+m'	0.52*	0.07	0.25*
		(0.27)	(0.1)	(0.13)
LB of Effort allocation effect: G4-G2	$\hat{l}$*	-0.14	0	-0.11
		(0.15)	(0.06)	(0.07)

***LB**: lower bound, **G1**: Improved seeds in the open RCT, **G2**: Local seeds in the open RCT, **G3**: Improved seeds in the double-blind RCT, **G4:** Local seeds in the double-blind RCT. The total effect of improved seeds is given by $\hat{g} + \hat{l} + \widehat{m}$ (G1- G2), the lower bound of the net effect by $\hat{g} + \widehat{m}^*$ (G3- G4), the lower bound of the behavioural effect by $\hat{l}' + \widehat{m}'$(G1- G3) and the lower bound of the effort allocation effect by $\hat{l}^*$ (G4- G2). Standard errors in parentheses, p-values of the t-test: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.*

We turn now to the results of the double-blind experiment. Following the model presented in section 3.3, they provide a lower bound of the net effect, i.e. a lower bound of the average increase in yield attributable to the higher productivity of the improved seeds: $\hat{g} + \widehat{m}^*$. In the North, the lower bound of the net effect on yield is 0.23 ton/ha (a 50% increase), significant at the 99% confidence level. In the East, it is 0.31 (a 38% increase), but it is not statistically different from zero. When both regions are pooled together, the average lower bound of the net effect is a 53% increase in yield, significant at the 99% confidence level.

By combining the results of the open and double-blind RCTs we obtain a lower and upper bounds of the net effect. At the aggregate level, the average net effect of improved seeds on yield is hence bounded between 0.29 ton per/ha to 0.44 ton/pa; in the North, between 0.23 ton/ha and 0.31/ton and in the East, between 0.31 ton/ha and 0.69 ton/ha. The difference between the open and double-blind RCTs is hence larger in the East, pointing toward a larger role of the behavioural response.

We now analyse the effect of the second treatment arm of the experiment, i.e. the change in the probability of receiving improved seeds. Following the model presented in section 3.3, the difference between group 4 and 2 gives a lower bound of the effort allocation effect, i.e. the increase in output caused by the increase in effort alone, $\hat{l}'$. We do not find any significant difference between group 4 and 2, which is consistent with the lack of behavioural response to a change in probability from 0% to 50% reported in Table 3.5. These results contrast with Bulte et al. (2014) where a significant effort allocation effect was found (or *pseudo-placebo* effect according to the terminology used in Bulte et al., 2014).

Following the model presented in section 3.3, the difference between group 1 and 3, provides a lower bound of the behavioural effect. It is large in the East: farmers being told that they have the improved seeds have yield 0.52 ton/ha higher than those who are told they have one chance out of two (significant at 90% confidence level). The behavioural effect represents close to 75% of the total effect in the East. Although it is not significant in the North, the contribution of the behavioural effect is large and significant when both regions are pooled together: 0.25 ton/ha or 57% of the total effect.

This highlights the importance of effort adjustment in order to reap the whole potential of the improved seeds. Farmers facing input constraints either because of limited access to land, fertilisers or man power will hence only benefit from a fraction of the potential increase in yield brought about by improved seeds.

We present in Table 3.8 the results of the regression analysis. Detailed results are shown in Table 3.10, Table 3.11 and Table 3.12 in the appendix. In model I, the only variables are the group dummies and, for the regression at the aggregate level, a dummy equal to one for the farmers in the North. In model II, we add controls related to farm management: labour (pre-harvest labour expressed in man days), a dummy for weeding (equal to 1 if the farmer conducted only one or no weeding), the size of the experimental plot (ha), the soil quality (a dummy equal to one for good soil), pesticide use (dummy), fertiliser use (dummy) and plot improvement works such as soil bunds or mulching (dummy). In model III, we add controls on production shocks, weather conditions and socioeconomic characteristics with the following set of variables: crop damage caused by pests, plant disease or fungi (dummy), extreme rainfall (dummy, self-reported), standardized precipitation index in level and square, a dummy for female-headed households, education of the household head (completed form IV, dummy), active role in the community such as seating in the village council (dummy) and risk loving attitude measured by a short risk preference experiment.[34] In model IV, we control for the extent of the social network and farmers' social interactions, i.e. the number of people the farmer can ask about farming related questions, the number of people the farmer can ask for help in cash, in kind or on the farm, a dummy

[34] Farmers were presented with a set of six lotteries ranked according to their riskiness (variance) and expected gain. Once the farmer had chosen his preferred lottery, a coin was flipped and he received the corresponding pay-off. The maximum amount was 4000 Tsh (circa 2 USD). Farmers who took one of the 2 riskiest lotteries were classified as risk lover (40% of the sample). See appendix for the protocol.

equal to one if the farmer consulted anybody about best farming practices, if he is a member of a social group (e.g. prayer group) or a member of a self-help group (e.g. credit union).

Table 3.8: Ordinary least squares regressions results

		Yield (ton/ha)			
	Model	**I**	**II**	**III**	**IV**
Both regions	Total Effect: G1-G2	0.46^{***}	0.45^{***}	0.44^{***}	0.43^{***}
		(0.14)	(0.14)	(0.15)	(0.14)
	LB of Net effect: G3-G4	0.25^{*}	0.25^{*}	0.21	0.21
		(0.12)	(0.13)	(0.13)	(0.14)
	LB of Behavioural effect: G1-G3	0.25^{*}	0.25	0.25	0.25^{*}
		(0.13)	(0.15)	(0.15)	(0.14)
	LB of Effort allocation effect: G4-G2	-0.06	-0.07	-0.07	-0.08
		(0.08)	(0.07)	(0.06)	(0.07)
	Observations	518	512	512	502
	R^2	0.169	0.201	0.284	0.300
North	Total Effect: G1-G2	0.31^{*}	0.28^{*}	0.24^{*}	0.24^{+}
		(0.08)	(0.08)	(0.08)	(0.09)
	LB of Net effect: G3-G4	0.23	0.26	0.24	0.22
		(0.12)	(0.15)	(0.15)	(0.19)
	LB of Behavioural effect: G1-G3	0.07	0.04	0.03	0.06
		(0.1)	(0.14)	(0.14)	(0.15)
	LB of Effort allocation effect: G4-G2	-0.00	-0.02	-0.03	-0.04
		(0.06)	(0.06)	(0.05)	(0.07)
	Observations	312	308	308	304
	R^2	0.072	0.129	0.173	0.253
East	Total Effect: G1-G2	0.69^{**}	0.67^{**}	0.65^{**}	0.64^{**}
		(0.26)	(0.25)	(0.25)	(0.25)
	LB of Net effect: G3-G4	0.31	0.21	0.14	0.24
		(0.31)	(0.32)	(0.29)	(0.3)
	LB of Behavioural effect: G1-G3	0.52^{*}	0.55^{**}	0.56^{**}	0.53^{**}
		(0.23)	(0.24)	(0.23)	(0.21)
	LB of Effort allocation effect: G4-G2	-0.14	-0.09	-0.06	-0.12
		(0.16)	(0.16)	(0.16)	(0.17)
	Observations	206	204	204	198
	R^2	0.085	0.131	0.257	0.282

***LB**: lower bound, **G1**: Improved seeds in the open RCT, **G2**: Local seeds in the open RCT, **G3**: Improved seeds in the double-blind RCT, **G4:** Local seeds in the double-blind RCT. The total effect of improved seeds is given by G1- G2 ($\hat{g} + \hat{l} + \hat{m}$), the lower bound of the net effect by G3-G4 ($\hat{g} + \hat{m}^*$), the lower bound of the behavioural effect by G1- G3 ($\hat{l}' + \hat{m}'$) and the lower bound of the effort allocation effect by G4- G2 ($\hat{l}^*$). Standard errors in parentheses (robust to clustering at the village level), p-values of the t-test: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.*

The regressions' results are consistent with the results of the comparison in means[35]. We do observe a slight decrease of the total effect and of the net effect once controls are added. This is likely due to the fact that part of the behavioural effect is controlled by the added variables. However, the low bound of the behavioural effect stays highly significant in the East and at the aggregate level, even

[35] The slight difference in the results for both regions pooled together between the regression and difference in mean analysis comes from the fact that in the former a dummy equal to one for farmers in the North and zero otherwise in added. The regression results conducted on each region separately do match as expected the results from the comparison in means.

after controlling for more than 20 variables (the p-value is just above 0.1 in model II and III at the aggregate level).

This suggests that a simple regression with a set of controls is not able to control appropriately for effort. A more complex model where the proxy variables for effort would be interacted with the improved seeds dummy could perhaps capture the increase in marginal productivity of effort. However, given the number of dimensions along which effort can be adjusted (e.g. quality, frequency), it might be challenging to properly account for the increased productivity. The combination of the double-blind and open RCT simplify greatly the estimation of various effect of interest: the total effect, the net effect, marginal productivity effect and the effort allocation effect.

3.6 Conclusion

While it is likely that no single factor can, on its own, explain the puzzle of the low adoption of improved seeds in Sub-Saharan Africa (see section 1.5 of the literature review), Suri (2011) showed that some hybrid maize seeds might not be profitable for a large part of the population, while Rosenzweig and Foster (2010) highlight the fact that many studies may overestimate the benefit of adoption because they neglect farmers' adjustment cost, i.e. they measure the total effect of adoption instead of the net effect.

We tested an improved maize seeds variety in Tanzania, the Situka M1, with a unique combination of open and double-blind randomized controlled trials (RCT). In the classic open RCT, by randomly allocating participants to a group receiving either the improved seeds or to another one receiving local seeds, the traditional issue of self-selection bias in studies based on observational data is bypassed and a direct and unbiased estimate of the total effect of the improved seeds is provided.

The advantage of combining an open and a double-blind RCTs is the randomization of the probability of receiving improved seeds. It allows us to disentangle the effect of the improved seeds *per se* from its effects which depends on a change in the management of the farm. The empirical contribution of this study is to show that this behavioural response plays a central role in driving the total effect of improved seeds. In the Eastern part of our sample, 75% of the total effect of improved seeds would not materialize without the behavioural response.

A second advantage of our design over classical open RCTs is to provide an upper and lower bounds estimate of the *net* effect of improved seeds, i.e. net of the adjustment in effort. At the aggregate level, the net effect of improved seeds is between 0.29 ton/ha and 0.44 ton/ha. Given this large increase in yield, the low adoption rate of improved seeds in the study area remains puzzling. A possible explanation is the large role of the behavioural response in driving the effect of improved seeds. Constraints on complementary inputs, such as low soil plot quality, might limit the scope of the behavioural response of some farmers. An interesting extension of the present paper would be to investigate the drivers of the behavioural effects as well as the distributional impacts of the improved seeds.

3.7 References

Alene, Arega D, et al. (2009), 'The economic and poverty impacts of maize research in West and Central Africa', *Agricultural Economics,* 40 (5), 535-50.

Amare, Mulubrhan, Asfaw, Solomon, and Shiferaw, Bekele (2012), 'Welfare impacts of maize–pigeonpea intensification in Tanzania', *Agricultural Economics,* 43 (1), 27-43.

Bank, World (2008), *World development report* (Washington DC).

Boisson, Sophie, et al. (2010), 'Field assessment of a novel household-based water filtration device: a randomised, placebo-controlled trial in the Democratic Republic of Congo', *PLoS One,* 5 (9), e12613.

Bulte, Erwin, et al. (2014), 'Behavioral responses and the impact of new agricultural technologies: evidence from a double-blind field experiment in Tanzania', *American Journal of Agricultural Economics,* 96 (3), 813-30.

Chassang, Sylvain, Snowberg, Erik, and Bowles, Cayley (2012), 'Accounting for Behavior in Treatment Effects: New Applications for Blind Trials', *Department of Economics, Princeton University*.

Diagne, Aliou and Demont, Matty (2007), 'Taking a new look at empirical models of adoption: average treatment effect estimation of adoption rates and their determinants', *Agricultural Economics,* 37 (2-3), 201-10.

Doss, Cheryl R (2003), *Understanding farm-level technology adoption: Lessons learned from CIMMYT's micro surveys in eastern africa* (CIMMYT).

Duflo, Esther, Kremer, Michael, and Robinson, Jonathan (2008), 'How high are rates of return to fertilizer? Evidence from field experiments in Kenya', *The American economic review*, 482-88.

Eskola, Elina (2005), *Agricultural marketing and supply chain management in Tanzania: A case study* (16: Economic and Social Research Foundation Dar es Salaam).

Evenson, Robert E and Gollin, Douglas (2003), 'Assessing the impact of the Green Revolution, 1960 to 2000', *Science,* 300 (5620), 758-62.

FADN, EU (2013), 'EU cereal farms report', in Directorate-General for Agriculture and Rural Development (ed.), (Brussell: European Commission).

Foster, Andrew D and Rosenzweig, Mark R (2010), 'Microeconomics of technology adoption', *Annual Review of Economics,* 2.

Giller, KE, et al. (2011), 'Communicating complexity: Integrated assessment of trade-offs concerning soil fertility management within African farming systems to support innovation and development', *Agricultural systems,* 104 (2), 191-203.

Glewwe, Paul, et al. (2004), 'Retrospective vs. prospective analyses of school inputs: the case of flip charts in Kenya', *Journal of development Economics,* 74 (1), 251-68.

Guttman, Nathaniel B (1999), 'Accepting the standardized precipitation index: a calculation algorithm', *Journal of the American Water Resources Association,* 35 (2), 311-22.

Kassie, Menale, Jaleta, Moti, and Mattei, Alessandra (2014), 'Evaluating the impact of improved maize varieties on food security in Rural Tanzania: Evidence from a continuous treatment approach', *Food Security,* 6 (2), 217-30.

List, John A (2011), 'Why economists should conduct field experiments and 14 tips for pulling one off', *The Journal of Economic Perspectives*, 3-15.

Magorokosho, C., Vivek, B., and MacRobertm, J. (2009), 'Characterization of maize germplasm grown in eastern and southern Africa: Results of the 2008 regional trials coordinated by CIMMYT', (Harare, Zimbabwe: CIMMYT).

McCarney, Rob, et al. (2007), 'The Hawthorne Effect: a randomised, controlled trial', *BMC medical research methodology,* 7 (1), 30.
Minot, N. (2010), 'Staple food prices in Tanzania', *Contributed Paper Prepared for the COMESA Policy Seminar Maputo*.
Rubin, Donald B (1974), 'Estimating causal effects of treatments in randomized and nonrandomized studies', *Journal of educational Psychology,* 66 (5), 688.
Shiferaw, Bekele, et al. (2014), 'Adoption of improved wheat varieties and impacts on household food security in Ethiopia', *Food Policy,* 44, 272-84.
Smale, Melinda, Byerlee, Derek, and Jayne, Thom (2011), 'Maize Revolutions in Sub-Saharan Africa', in World Bank Policy Research working paper (ed.).
Suri, Tavneet (2011), 'Selection and comparative advantage in technology adoption', *Econometrica,* 79 (1), 159-209.
Thomas, Duncan, et al. (2004), 'Causal Effect of Health on Labor Market Outcomes: Evidence from a Random Assignment Iron Supplementation Intervention'.
Tumbo, SD, et al. (2012), 'Impact of projected climate change on agricultural production in semi-arid areas of Tanzania: a case of same district', *African Crop Science Journal,* 20 (2), 453-63.

3.8 Appendix A: Further results

Table 3.9: Did randomization work?

	Averages				P-values of the t-test					
	G1	G2	G3	G4	1=2	1=3	1=4	2=3	2=4	3=4
Land owned fully (ha)	2.83	3.01	2.64	2.62	0.53	0.55	0.48	0.25	0.21	0.95
Land owned (ha)	2.96	3.06	2.76	2.69	0.74	0.52	0.37	0.36	0.25	0.83
Land farmed (ha)	1.46	1.50	1.43	1.38	0.73	0.81	0.52	0.58	0.33	0.72
Oxen (d.)	0.23	0.22	0.30	0.29	0.89	0.19	0.29	0.14	0.22	0.81
Rich (self-reported, d.)	0.26	0.26	0.29	0.23	0.96	0.55	0.61	0.57	0.57	0.29
Female headed household	0.09	0.13	0.15	0.12	0.29	0.17	0.51	0.70	0.73	0.50
Age household head	44.6	45.5	48.4	47.3	0.51	0.02	0.07	0.06	0.21	0.52
Dependency ratio (no of dependent per adult)	0.53	0.55	0.58	0.42	0.70	0.51	0.16	0.73	0.06	0.05
Household size	5.14	5.03	5.10	5.16	0.67	0.90	0.93	0.78	0.61	0.83
Education head (years of education)	6.01	5.89	6.02	5.88	0.76	0.99	0.76	0.75	0.98	0.74
Education: one household member up to form IV	0.45	0.47	0.52	0.42	0.67	0.22	0.68	0.40	0.40	0.12
Education: tow household member up to form IV	0.22	0.21	0.13	0.15	0.91	0.07	0.17	0.09	0.19	0.68
Social network (number of people)	1.82	1.89	2.13	1.96	0.71	0.23	0.55	0.35	0.78	0.56
Social network (number of relatives)	2.27	2.49	2.76	2.47	0.41	0.13	0.52	0.35	0.95	0.40
Member of a social association (e.g. youth, women)	0.78	0.75	0.69	0.72	0.54	0.12	0.30	0.33	0.64	0.63
Member of a self-help group (e.g. Sacco, vicoba)	0.40	0.39	0.39	0.33	0.81	0.86	0.27	0.96	0.36	0.37
Active role in the community	0.28	0.24	0.26	0.28	0.39	0.72	0.91	0.65	0.48	0.82
Oxen (d.)	0.23	0.22	0.30	0.29	0.89	0.19	0.29	0.14	0.22	0.81
Rich (self-reported, d.)	0.26	0.26	0.29	0.23	0.96	0.55	0.61	0.57	0.57	0.29
Observations	148	166	121	126						

Table 3.10: Ordinary Least Squares: Both regions

Dependent variable: Harvest (kg) Base category: local seeds in the open RCT	I	II	III	IV
Improved seeds, open RCT	0.46***	0.45***	0.44***	0.43***
	(0.14)	(0.14)	(0.15)	(0.14)
Improved seeds, double-blind RCT	0.21+	0.20	0.18	0.18
	(0.13)	(0.14)	(0.15)	(0.15)
Local seeds, double-blind RCT	-0.04	-0.05	-0.03	-0.03
	(0.07)	(0.06)	(0.06)	(0.07)
North	-0.57**	-0.54**	-1.01***	-1.06***
	(0.20)	(0.20)	(0.24)	(0.25)
Labour (man day, pre-harvest)		0.00	0.01	0.01
		(0.01)	(0.01)	(0.01)
One or no weeding		-0.29**	-0.20**	-0.21**
		(0.12)	(0.08)	(0.09)
Experimental plot size (ha)		0.18	-0.01	0.01
		(0.58)	(0.52)	(0.56)
Good soil (dummy)		0.12	0.12+	0.11
		(0.08)	(0.07)	(0.07)
Pesticide (dummy)		-0.23***	-0.20**	-0.20***
		(0.07)	(0.07)	(0.06)
Fertiliser (dummy, mostly manure)		-0.06	-0.02	-0.01
		(0.07)	(0.08)	(0.08)
Plot improvement (e.g. soil bund,		0.04	0.06	0.06
Mulching, dummy)		(0.15)	(0.13)	(0.13)
Standardized precipitation Index			0.72**	0.74**
(peak rainfall month, ARC 2 dataset)			(0.27)	(0.29)
Standardized precipitation Index^2			-1.36	-1.31
			(0.93)	(0.94)
Crop damage due to pest, diseases,			-0.17+	-0.17
or fungi (dummy)			(0.11)	(0.11)
Extreme precipitation (self-reported,			-0.13*	-0.14*
dummy)			(0.07)	(0.08)
Female headed household (dummy)				-0.03
				(0.12)
Basic education (Standard IV,				0.01
dummy)				(0.07)
Active role in the community				0.16*
(dummy)				(0.08)
Dependency ratio				0.00
(no of dependent per adult)				(0.10)
Risk lover (risk experiment, dummy)				-0.02
				(0.05)
Information network (number of				0.00
people the farmer can ask farming related				(0.00)
question)				
Help network (number of people the				0.01
farmer can ask help in cash, in kind or on the				(0.01)
farm)				
Consulted a farmer or an extension				0.04
agent on farming practices (dummy)				(0.06)
Member of a social association				-0.06
(e.g. prayer groups , dummy)				(0.12)

Member of a self-help group (e.g. credit union, dummy)				-0.10 (0.09)
Constant	1.00***	0.99***	1.46***	1.46***
	(0.15)	(0.18)	(0.25)	(0.25)
Observations	518	512	512	502
R^2	0.169	0.201	0.284	0.300

Standard errors in parentheses (robust to clustering at the village level) * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 3.11: Ordinary Least Squares: North

Dependent variable: Harvest (kg) Base category: local seeds in the open RCT	I	II	III	IV
Improved seeds, open RCT	0.31^{*}	0.28^{*}	0.24^{*}	0.24^{+}
	(0.08)	(0.08)	(0.08)	(0.09)
Improved seeds, double-blind RCT	0.23	0.23	0.20	0.19
	(0.13)	(0.15)	(0.15)	(0.21)
Local seeds, double-blind RCT	-0.00	-0.02	-0.03	-0.04
	(0.06)	(0.06)	(0.05)	(0.07)
Labour (man day, pre-harvest)		0.01	0.01	0.01
		(0.01)	(0.01)	(0.01)
One or no weeding		-0.22^{*}	-0.22^{*}	-0.20^{+}
		(0.06)	(0.06)	(0.08)
Experimental plot size (ha)		-0.06	-0.03	-0.06
		(0.63)	(0.63)	(0.66)
Good soil (dummy)		0.03	0.05	0.09^{**}
		(0.04)	(0.04)	(0.01)
Pesticide (dummy)		-0.20^{+}	-0.16	-0.15
		(0.09)	(0.11)	(0.09)
Fertiliser (dummy, mostly manure)		-0.09^{*}	-0.10	-0.11^{**}
		(0.02)	(0.04)	(0.02)
Plot improvement (e.g. soil bund,		0.13^{**}	0.13^{**}	0.11^{*}
Mulching, dummy)		(0.03)	(0.03)	(0.03)
Standardized precipitation Index			11.77^{+}	7.94^{**}
(peak rainfall month, ARC 2 dataset)			(4.85)	(0.83)
Standardized precipitation Index^{2}			-16.53	-10.86^{**}
			(7.32)	(1.31)
Crop damage due to pest, diseases,			-0.22^{*}	-0.18
or fungi (dummy)			(0.06)	(0.12)
Extreme precipitation (self-reported,			-0.11^{**}	-0.14^{*}
dummy)			(0.02)	(0.04)
Female headed household (dummy)				-0.09
				(0.14)
Basic education (Standard IV,				0.06
dummy)				(0.05)
Active role in the community				0.17
(dummy)				(0.09)
Dependency ratio				-0.19^{***}
(no of dependent per adult)				(0.01)
Risk lover (risk experiment, dummy)				-0.07
				(0.06)
Information network (number of				-0.01
people the farmer can ask farming related question)				(0.00)
Help network (number of people the				0.01^{**}
farmer can ask help in cash, in kind or on the farm)				(0.00)
Consulted a farmer or an extension				0.09
agent on farming practices (dummy)				(0.06)
Member of a social association				-0.06
(e.g. prayer groups , dummy)				(0.12)
Member of a self-help group				-0.01
(e.g. credit union, dummy)				(0.09)
Constant	0.46^{**}	0.48^{*}	-1.35	-0.74^{***}

	(0.08)	(0.13)	(0.59)	(0.04)
Observations	312	308	308	304
R^2	0.072	0.129	0.173	0.253

Standard errors in parentheses (robust to clustering at the village level) * *$p < 0.10$, ** $p < 0.05$, *** $p < 0.01$*

Table 3.12: Ordinary Least Squares: East

Dependent variable: Harvest (kg) Base category: local seeds in the open RCT	I	II	III	IV
Improved seeds, open RCT	0.69**	0.67**	0.65**	0.64**
	(0.26)	(0.25)	(0.25)	(0.25)
Improved seeds, double-blind RCT	0.17	0.12	0.09	0.11
	(0.27)	(0.28)	(0.27)	(0.26)
Local seeds, double-blind RCT	-0.14	-0.09	-0.06	-0.12
	(0.16)	(0.16)	(0.16)	(0.17)
Labour (man day, pre-harvest)		-0.01	0.01	0.01
		(0.02)	(0.02)	(0.02)
One or no weeding		-0.39+	-0.22	-0.23
		(0.25)	(0.17)	(0.20)
Experimental plot size (ha)		1.37	0.20	1.32
		(2.35)	(1.65)	(1.70)
Good soil (dummy)		0.17	0.12	0.13
		(0.16)	(0.15)	(0.14)
Pesticide (dummy)		-0.67+	-0.62*	-0.54
		(0.38)	(0.34)	(0.36)
Fertiliser (dummy, mostly manure)		0.20	0.51	0.52
		(0.51)	(0.43)	(0.39)
Plot improvement (e.g. soil bund,		-0.12	-0.11	-0.07
Mulching, dummy)		(0.43)	(0.35)	(0.33)
Standardized precipitation Index			0.56*	0.57*
(peak rainfall month, ARC 2 dataset)			(0.27)	(0.29)
Standardized precipitation Index^2			-2.51	-2.35
			(1.80)	(1.90)
Crop damage due to pest, diseases,			-0.01	0.05
or fungi (dummy)			(0.27)	(0.33)
Extreme precipitation (self-reported,			-0.11	-0.09
dummy)			(0.25)	(0.26)
Female headed household (dummy)				0.15
				(0.18)
Basic education (Standard IV,				-0.09
dummy)				(0.12)
Active role in the community				0.11
(dummy)				(0.12)
Dependency ratio				0.10
(no of dependent per adult)				(0.13)
Risk lover (risk experiment, dummy)				0.01
				(0.12)
Information network (number of				0.01
people the farmer can ask farming related				(0.01)
question)				
Help network (number of people the				0.01
farmer can ask help in cash, in kind or on the				(0.02)
farm)				
Consulted a farmer or an extension				-0.05
agent on farming practices (dummy)				(0.13)
Member of a social association				0.00
(e.g. prayer groups , dummy)				(.)
Member of a self-help group				-0.24
(e.g. credit union, dummy)				(0.17)
Constant	0.96***	0.93**	1.64***	1.48**

	(0.13)	(0.39)	(0.44)	(0.60)
Observations	206	204	204	198
R^2	0.085	0.131	0.257	0.282

Standard errors in parentheses (robust to clustering at the village level) * $p < 0.10$, ** $p < 0.05$, *** $p < 0.$

3.9 Appendix B : Risk preference experiment

The respondent is asked to choose between the different farming plots (plot 1 to plot 6); Each plot gives either the bad harvest yield or a good harvest yield. For instance, plot 2 gives 900 Shillings if the season is bad (bad harvest), but it gives 1800 Shillings if the season is good (good harvest).

Extentionist, please read the following: Imagine you can select 1 of 6 plots. On plot one, you earn 1000 Tsh if the season is bad (HEAD) and also 1000 Tsh if the season is good (TAIL); on plot two 900 Tsh if the season is bad or 1800 Tsh if the season is good; on plot three 800 Tsh or 2400 Tsh; on plot four 600 Tsh or 3000 Tsh; on plot five 200 or 3600 Tsh and on plot six 0 or 4000. In each plot, there is a one chance in two to get the bad and good harvest, that is: a good season is as likely as a bad season. Please, take a moment to compare the six different plots and then tell me which plot is the best for you.

Extentionist: show the boxes below to the farmers and explain him again how it works.

Plot 1.___	Plot 2.___	Plot 3.___	Plot 4.___	Plot 5.___	Plot 6.___
Bad harvest (Head) 1000	Bad harvest (Head) 900	Bad harvest (Head) 800	Bad harvest (Head) 600	Bad harvest (Head) 200	Bad harvest (Head) 0
Good harvest (Tail) 1000	Good harvest (Tail) 1800	Good harvest (Tail) 2400	Good harvest (Tail) 3000	Good harvest (Tail) 3600	Good harvest (Tail) 4000

4 Avoiding the 'Family Tax': Social pressure and hiding in village economies

Xavier Vollenweider[1] and Salvatore Di Falco[2]

Abstract

Based on the field experiment on maize seeds conducted in Tanzania and presented in chapter 3, we test the hypothesis that individuals try to escape forced solidarity when facing favourable conditions. We find that farmers who were allocated the improved seeds decrease the number of their social interactions, particularly if they have a large number of relatives in the village. We interpret these results as an evidence that farmers who were assigned the improved seeds adopted an evasive behaviour to escape the redistributive pressure from their social network. Furthermore, it suggests that the pressure to share increases with the size of the social network.

[1] Department of Geography and Environment, London School of Economics and Political Science, United Kingdom. Contact author: Xavier Vollenweider: X.Y.Vollenweider@lse.ac.uk.

[2] Department of Economics, University of Geneva, Switzerland.

4.1 Introduction

Social networks play an important role in the livelihood of rural communities in developing countries by providing informal insurance and credit when markets are imperfect or absent (Anderson and Baland 2002; Coate and Ravallion 1993; Fafchamps 1992; Fafchamps and Lund 2003; Ligon et al. 2002; Rosenzweig 1988; Townsend 1994; Udry 1994). Households' expectations of future assistance and transfers are key motivations behind participation in these networks. Other explanations may, however, apply. Altruism, guilt, social pressure to share resources and potential social sanctions also seem to play a crucial role in shaping individual behaviour in networks (Alger and Weibull 2010; Barr and Stein 2008; Foster and Rosenzweig 2001; Leider et al. 2009; Ligon 2011; Ligon and Schechter 2012; Platteau 2000). Social relations define obligations for the network's members.[3] The more successful members of the network must help the least successful or unlucky members of the social network.

Recently, some observational and experimental evidence has been provided indicating that these obligations may trigger an *evasive* response. Households anticipating that their future income will be 'taxed' by kin and neighbours may alter their consumption and investment decisions. They may, for instance, try to escape these obligations by spending more on non-sharable goods and keeping less liquidity (Di Falco and Bulte 2011). Individuals may attempt to fend off network requests by 'pretending to be poor' (Baland et al. 2011) or by concealing their assets and making more investments when not observed by kinship members (Jakiela and Ozier 2016). In this paper, we directly analyse the role of hiding as an evasive response to social network pressure by using a field experiment in rural Tanzania.[4] We randomly assigned a positive income shock to some farmers by allocating improved seeds to some farm households, while others were assigned a traditional low yielding variety. The expected future income of households with improved seeds is therefore raised. We found that individuals who were assigned the improved seeds reduced their social interactions if they counted a large number of relatives within the village.

In section 4.2, we introduce the experimental design, the main idea behind the study, the variable used as proxy of the social interactions and social network. In section 4.3, we introduce a simple model to further

[3] In this respect, Scott (1976) and Platteau (1991) refer to the 'moral economy'.
[4] This is the same experiment used in chapter 3 with the slight difference that we use only the results from the open randomized controlled trial, i.e. the farmers who were told which type of seeds they received (i.e. group 1 and 2 in chapter 3).

the empirical analysis. In section 4.4, we present the econometric strategy, in section 4.5, we present the results and we conclude in section 4.6.

4.2 Data and Design of the Field Experiment

The field experiment was conducted in two areas of Tanzania, the South East (Morogoro) and the North (Karatu) with a sample of 320 farmers.[5] Improved (high yielding) maize seeds were allocated randomly to half of the sample in 2013. The remaining half received traditional (low yielding) varieties. Maize is a key crop for these areas and is basically grown by all farmers mostly for their own consumption. In good years or among large farms, surpluses are marketed. The improved variety of Situka-M1 was released in 2001 by the Selian Agricultural Research Institute (SARI). It has a very high yield potential of 3-5 ton/ha and its optimal production altitude range is 1000-1500 m. The traditional variety has a potential of 0.5-1 ton/ha under similar conditions. In Tanzania, the improved variety is grown in the South Eastern and Northern regions where our study districts are located. It is the second most important open pollinated variety (OPV) following the Staha variety grown in our study areas.[6] The Situka-M1 is hence well-known and considered as a high yielding variety by farmers. Thus, allocating Situka-M1 constituted a positive shock on expected income as shown by the results of chapter 3 where harvest was on average 50% higher with the Situka-M1 than the local variety of seeds.

The goal of the present chapter is to compare social interactions between the farmers who got the improved seeds and the control group's farmers who got local seeds. Each farmer was asked in the end-line survey the number of social interactions they have had over the study period according to seven categories summarized in Table 4.1. Let us for the moment focus on asking for help on the farm. Asking for help on the farm is a very common social interaction in the village. In this way farmers can use extra units of labour; this is a direct and tangible benefit of a social network. Asking for help entails one important implication: visibility. The people giving a helping hand can guess the farmer's future harvest. If it is a bumper crop, the farmer exposes himself to solidarity requests from those that helped him, and from less fortunate members of the community having heard of his bumper crop. Hence there is a choice

[5] We use the same experiment as in chapter 3, using here only the data from farmers who knew what type of seeds they were given (i.e. group 1 and 2 in the chapter 3) and not the double-blind experiment. Indeed, we want to observe the participants' behaviour when they know that they have a positive income shock.
[6] About 12% of farmers used Situka-M1 during the 2010/11crop calendar. The variety is tolerant to drought, and maize streak and grey leaf spot diseases.

between, on the one side, asking for help and being exposed to a solidarity tax and, on the other side, not benefiting from help and hiding from a solidarity request. A farmer having received improved seeds will think twice before asking for help: in the eyes of the community, he might become one of the lucky few to whom one turns to in case of hardship.

The social interactions recorded in the survey vary in terms of the visibility they entail. Discussing the type of seeds received in the experiment and asking for help on the experimental plot or, more generally, on the farm implies a high visibility. By contrast, asking for help in cash or in kind or for information about best practices or output and land market does not require revealing the type of seeds received in the experiment (low visibility).

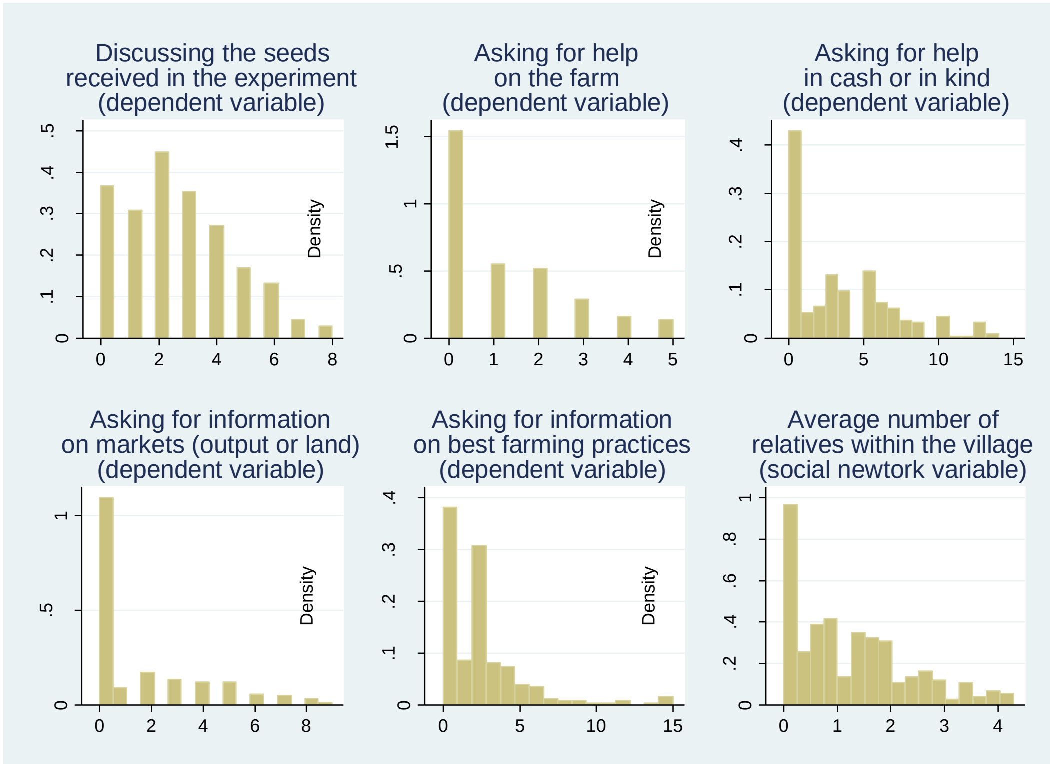
Table 4.1: Social interactions analysed

	Visibility	Improved	Local	Diff.
Discussing the seeds received in the experiment (count)	*high*	3.33	4.05	-0.72
		(3.70)	(5.42)	(0.53)
Asking for help on the farm (count)	*high*	1.74	2.06	-0.32
		(2.64)	(2.87)	(0.31)
Asking for help on the experimental plot (harvest) (d)	*high*	0.10	0.10	-0.00
		(0.30)	(0.30)	(0.03)
Asking for help on the experimental plot (pre-harvest) (d)	*high*	0.36	0.33	0.03
		(0.48)	(0.47)	(0.05)
Asking for help in cash or in kind (count)	*low*	4.19	4.23	-0.04
		(6.02)	(6.58)	(0.71)
Asking for information on best farming practices (count)	*low*	2.41	2.57	-0.15
		(3.45)	(2.91)	(0.36)
Asking for information on markets (output or land) (count)	*low*	2.89	2.31	0.58
		(6.03)	(3.52)	(0.55)

Standard deviations and standard errors in parentheses, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$, (d) stands for dummy variables, (count) stands for the number of people with whom each farmer had one of the social interaction.

Summary statistics on the social interactions variables are presented in Table 4.1. We see that the number of people each farmer interacts with oscillates between 1 and 3. Furthermore, while asking for help on the experimental plot is relatively common (more than one third of the sample does it), asking for help at harvest time is much less common (10 percent). There are no statistically significant differences between farmers in the improved and local seeds category. However, the difference in the propensity to ask may depend on the size of the social network. Indeed, visibility may come at a greater cost with a large network because larger networks may imply larger pressure for sharing. In section 4.3 we will present a formalisation of this argument based on expected utility theory. The distributions of the count data variables are shown in Figure 4.1.

Figure 4.1: Distribution of the main variables



The social network variable is based on the number of people each farmer could ask for help or information. As each category has been recorded separately, it is likely that some networks overlap (e.g. friends ready to give information on best farming practices are ready to give information on output and land markets). We have chosen to express the network size as the average size of all social network categories in order to reduce the impact of double-counting and capture more closely the number of people each farmer knows.[7] The social network is measured in three layers: people in general, relatives inside the village, and relatives outside the village. We opted for relatives inside the village as it is in this layer that solidarity should be the strongest as geographical proximity and kinship are the major determinants of risk-sharing network formation (Fafchamps and Gubert 2007). Farmers can ask for help or information on average from 1.5 relatives in the village. Its distribution is shown in Figure 4.1.

[7] An alternative strategy is to express the network size as the sum of all social network categories in order to capture the *intensity* or the *usefulness* of each member of a farmer's social network: a friend who can help with cash, kind or help on the farm is worth more than a friend who can only provide help on the farm. As we found very similar results, we therefore present results only in terms of averaged networks.

We also controlled for external drivers which could trigger the solidarity within the villages such as pest damage and excessive rainfall. Field visits revealed heavy damage to some fields from termites in the South East and from fungi and excessive rainfall in the North, close to one quarter of the farmers reported crop damage due to pests. The 2013 season was rather wet as half of the farmers reported rain above the normal (70% in the North). This is also recorded with satellite observations from the African Rainfall Climatology Version 2 dataset (ARC2, Novella and Thiaw 2013): the standardized precipitation index for March was 0.64 in the North with a maximum of 0.91 in one village.

We also tried to limit the effect of omitted variable bias by controlling for socioeconomic characteristics possibly correlated with social interactions such as belonging to a self-help group or other economic association (e.g. primary society, a rotating savings and credit association or funeral society), for belonging to a social group (e.g. a prayer group or a youth organisation), the size of the household, the education, female headed household, the age of the household head, the size of the whole farm measure in ha, oxen holding, and for living in the North. Lastly, a farmer with improved seeds sowed on a highly visible plot (e.g. near the homestead) is more likely to be identified by his peers as expecting a high harvest. This could trigger demand for help from his social network even in the absence of a request to help or information from his part. In order to control for the visibility of the plot, we add a variable measuring the walking distance from the homestead to the experimental plot (measured in minutes). A more remote plot is less observable. Summary statistics are displayed in Table 4.2. Note that the average size of the farm is just above 1.5 ha, i.e. most farmers are small-scale subsistence with no surplus for sale.

Table 4.2: Explicative variables

	Mean	Standard Deviation
Improved seeds(d)	0.47	0.50
Number of relatives within the village	1.50	1.57
North(d)	0.59	0.49
Plot size (ha)	0.11	0.04
Distance to plot (minutes)	18.86	19.12
Farm size (ha)	1.58	1.23
Oxen (d)	0.23	0.42
Labour (man day)	9.27	6.87
Pest damage (d)	0.23	0.42
Standardized Precipitation Index	0.22	0.66
Female headed household (d)	0.11	0.32
Age household head(d)	45.74	12.43
Household size	5.11	2.24
Secondary education (d)	0.60	0.49

Risk averse(d)[8]	0.22	0.41
Leadership role in the community(d)	0.31	0.56
Member of a self-help group (d)	0.39	0.49
Member of a religious association (d)	0.76	0.43

(d) stands for dummy variables.

Despite the low level of attrition (around 9%), we checked that attrition was not correlated with geo-physical characteristics. We could find no significant association. Furthermore, we compared farmers in each group according to a set of demographic, socioeconomic and geophysical characteristics and we could not find any statistical difference (Table 4.3). It hence appears that the randomization has worked and that the groups of treated and untreated farmers are comparable.[9]

Table 4.3: Randomization results: two comparable groups of farmers

	Improved seeds	**Local seeds**	**p-value of a t-test of equality of mean**	**N**
Number of relatives within the village	1.57	1.43	0.43	313
Farm size (ha)	1.48	1.42	0.6	286
Distance to plot (minutes)	17.68	18.84	0.528	300
Standardized precipitation index: March	0.18	0.27	0.209	314
Dependency ratio	0.29	0.27	0.402	304
Age of the household head	45.41	44.42	0.377	286
Household size	4.86	5.08	0.302	295
Motorbike (d)	0.17	0.14	0.336	314
Oxen (d)	0.22	0.23	0.885	314
Rich (self-reported) (d)	0.26	0.26	0.963	314
Member of a religious association (d)	0.75	0.78	0.535	314
Member of a self-help group (d)	0.39	0.4	0.813	314
Leadership role in the community (d)	0.24	0.28	0.39	314
No education(d)	0.41	0.39	0.75	314
Female (d)	0.13	0.09	0.294	314
Pest damage (d)	0.27	0.19	0.135	314

[8] Farmers were presented with a set of six lotteries ranked according to their riskiness (variance) and expected gain. Once the farmer had chosen his preferred lottery, a coin was flipped and he received the corresponding pay-off. The maximum amount was 4000 Tsh (circa 2 USD). Farmers who took one of the 2 riskiest lotteries were classified as risk lover seeking (40% of the sample). See appendix B of chapter 3 for the protocol.

[9] We cannot be sure however that the randomization worked in terms of the longer history of giving and receiving help. Nevertheless, as farmers do not differ either in terms of their socio-demographic characteristics, or social capital as measured by the dummy variable equal to one if the farmer has a leadership role in the community, by the dummy variable equal to one if he is member of a self-help group or a religious association, or by the social network size, we assume that the randomization worked in terms of the longer history of help.

(d) stands for dummy variables.

4.3 Conceptual background

In order to define the notion of evasive behaviour, we present a simple model of the decision to ask for help on the farm when confronted to a positive income shock. In our experiment, this positive shock is represented by receiving the improved seeds. The model serves as a general motivational device for the empirical work. We do not attempt to identify the structural parameters of a behavioural model which would be estimated and tested in the empirical section.

We start with a model where social network is only given a negative role: the larger it is, the larger is the probability of being asked for help. Positive effects of social networks, such as their role as a channel of financial resources and information, are then discussed. We comment at the end of the section the implications of considering other types of social interactions than asking for help on the farm.

The model has three periods. In the first period, the farmer can ask for help on the farm. In the second period, he draws utility from the consumption of the harvest and saves in the form of an asset the potential additional income he made from it In the third period, a member of his social network can ask him for help, hereafter referred to as a tax. The farmer pursues an asset-smoothing strategy rather than a consumption-smoothing one. Therefore, if he is confronted with a tax in the third period, he pays it by decreasing consumption rather than drawing down on the savings. Lastly, the farmer does not discount future utility and does not consider the utility of saving when deciding about asking for help on the farm, he takes into account only the marginal utility of consumption.

The utility of asking and not asking for help on the farm is given respectively:

$$\mathrm{V^{Ask}} = \mathrm{U(C_H + b) + \pi_A U(c_F - d) + (1 - \pi_A) U(c_F)} \quad (1)$$

$$\mathrm{V^{Not}} = \mathrm{U(C_H) + \pi_N U(c_F - d) + (1 - \pi_N) U(c_F)} \quad (2)$$

where $U(\ )$ is a utility function, C_H is the consumption at harvest time (i.e. at the second period), b is the added consumption at harvest time due to having asked for help on the farm (e.g. thanks to better land preparation, weeding, slashing etc.), π_A is the probability of being taxed in the third period if one has asked for help in the first period, π_N is the probability of being taxed in the third period if one has *not* asked for help in the first period, C_F is the consumption level in the third period and d is the tax. Both C_H and C_F in period 2 and 3 are assumed to be constant, the only source of risk comes from the probability of being taxed.

We assume that improved seeds increases harvest in period 2. This rests on the observation that, on average, improved seeds increase yields by more than 60% following the results presented in chapter 3. Improved seeds might however not always be adapted to the needs of small scale farmers as, for instance, it has been found to be the case of hybrid maize seeds in Kenya (Suri 2011). Furthermore, we showed in chapter 3 that a large part of the increase in yield was driven by the farm management. Therefore, differences in know-how, plot quality or inputs access could imply that some farmers derive only a low benefit from the improved seeds. The Situka-M1, the improved seeds we distributed, is however well perceived by farmers who took part in the experiment according to discussions with key informants. Furthermore, as it is an open-pollinated variety and not a hybrid variety, it does not require large application of costly inorganic fertilisers to increase yield. We therefore assume that receiving improved seeds is a positive shock.

Furthermore, we assume that the productivity of labour with improved seeds is higher than with local seeds so that $\mathrm{b}_I > \mathrm{b}_L$, where the subscripts I and L denote improved and local seeds respectively. The maximum consumption in period 2 is given by $\mathrm{C_H} + \mathrm{b}_I + z$. Any additional harvest is sold on the market and the resulting income is saved.

Four additional assumptions are made:

1. Farmers' harvest is not known precisely by the members of their social network;
2. Asking for help on the farm implies revealing one's expected harvest to the members of their social network;
3. The probability of being taxed, $\pi\ (n, k, f)$, increases with the size of the network, n, the wealth reputation, k and the *help debt*, f:

$$\frac{\partial \pi(n,k,f)}{\partial i} > 0 \tag{3}$$

   where i is either $k, n,$ or f. The size of the social network, n, is the number of people the farmer has to help in case they ask. The wealth reputation is the amount of saving that the members of the social network think the farmer has. A larger harvest increases the wealth reputation. The help debt, f, is the number of times the farmer has asked for help minus the number of times he has been asked for help in the past.

4. The increase in the probability of being taxed, when the wealth reputation increases, is higher when the size of the social network is larger or when the help debt is larger. More generally:

$$\frac{\partial^2 \pi(n,k,f)}{\partial k \partial j} > 0 \tag{4}$$

where i, j, is k, n, or f and $i \neq j$.

Assumption (1) rests on the observation that in the developing world 'people consciously try to decrease observability of their income and wealth' by avoiding the disclosure of any information 'surrounding grain storage, livestock, and other assets to their counterparts' (Fafchamps 1992). In the model below, it implies that the expected harvest is not known by the other members of the network.

Assumption (2) is based on the fact that 'it is easy for an experienced farmer to guess crop yield by observing standing crops at harvest' (idem).

Assumption (3) relies on the observations that wealthy members of a social network have a moral duty to help the poor and unlucky ones. An increase in the wealth reputation should therefore increases the risk of being taxed. Furthermore, people with a large network often play a central role in the village's community life and it is toward them that one turns for assistance when confronted with a shock.[10] Lastly, by asking for help, one contracts a debt which will need to be paid pack. There is hence a higher likelihood to be taxed if one asked for help in the first period.

Assumption (4) implies that an increase in wealth reputation increases more the risk of a tax when the network is larger because the number of people who can tax is larger. The same idea holds for the help debt: if someone has a long history of relying on the solidarity of his network, i.e. his help debt his large, then we can expect that as soon has his wealth reputation increases, his creditors will come to ask him their due. The more numerous they are, the higher is the chance that he will be confronted by a tax.

---

[10] Another way of defending assumption (3) is to restrict the definition of an 'increase' in *n* as an increase in the size of the ego-network by the addition of a new member not connected to the previous members of the ego-network. A larger network would hence stretch over smaller unconnected ego-networks. Let us imagine a social network composed of 3 individuals forming a triangle with each individual represented as a dot in one corner. Let one of the three individuals meet two other people and form a separate network with them. Graphically, there are now two triangles head to head. The individual who is at the junction can be taxed by 4 people. By contrast, each member of his ego-network can be taxed by only two people. Hence, when an ego network increases by unconnected members of the ego network, the risk of being taxed increases.

The marginal utility of asking for help on the farm is given by:

$$\Delta V = \mathrm{U}(\mathrm{C_H} + \mathrm{b}) - \mathrm{U}(\mathrm{C_H}) + \pi_\mathrm{A}\mathrm{U}(\mathrm{c_F} - \mathrm{d}) + (1 - \pi_\mathrm{A})\mathrm{U}(\mathrm{c_F}) - \pi_\mathrm{N}\mathrm{U}(\mathrm{c_F} - \mathrm{d}) - (1 - \pi_\mathrm{N})\mathrm{U}(\mathrm{c_F}) \tag{5}$$

$$\Delta V = \mathrm{U}(\mathrm{C_H} + \mathrm{b}) - \mathrm{U}(\mathrm{C_H}) + (\pi_\mathrm{A} - \pi_\mathrm{A})[\mathrm{U}(\mathrm{c_F} - \mathrm{d}) - \mathrm{U}(\mathrm{c_F})] \tag{6}$$

$$\Delta V = \mathrm{U}(\mathrm{C_H} + \mathrm{b}) - \mathrm{U}(\mathrm{C_H}) - (\pi_\mathrm{A} - \pi_\mathrm{N})[\mathrm{U}(\mathrm{c_F}) - \mathrm{U}(\mathrm{c_F} - \mathrm{d})] \tag{7}$$

For notational convenience, we assume that b and d imply only a marginal change in consumption. We hence have:

$$\Delta V \approx U'(C_H)b - (\pi_\mathrm{A} - \pi_\mathrm{N})U'(c_F - d)d \tag{8}$$

The marginal cost, MC, and the marginal benefit, MB, can hence be expressed as:

$$MB = U'(C_H)b \tag{9}$$

$$MC = (\pi_\mathrm{A} - \pi_\mathrm{N})U'(c_F - d)d \tag{10}$$

The MB derives from the marginal increase in consumption brought about by the helping hand, the marginal cost derives from the increase in the probability of being taxed. The farmer does not ask for help if the MC is larger than the MB:

$$\Delta V \approx MB - MC < 0 \tag{11}$$

Equation (11) is the decision to hide: it results from the willingness to sacrifice MB in order to decrease the risk of being taxed.

Now, let us investigate the MB and the MC with improved and local seeds. With local seeds, the *MB* and MC are given by:

$$MB_L = U'(C_H)b_L \tag{12}$$

$$\mathrm{MC}_L = \frac{\partial \pi(n,k,f)}{\partial f} U'(c_F - d)d \tag{13}$$

where the subscript L is used to denote local seeds. With improved seeds, the MC and MB are given by:

$$MB_I = U'(C_H + z)b_I \tag{14}$$

$$\mathrm{MC}_I = \left(\frac{\partial \pi(n,k,f)}{\partial f} + \frac{\partial \pi(n,k,f)}{\partial k}\right) U'(c_F - d)d \tag{15}$$

where the subscript I is used to denote improved seeds The effect of improved seeds on MB and MC is hence given by:

$$\Delta MB = MB_I - MB_L = U'(C_H + z)b_I - U'(C_H)b_L \lesseqqgtr 0 \quad (16)$$

$$\Delta MC = MC_I - MC_L = \frac{\partial \pi(n,k,f)}{\partial k} U'(c_F - d)d > 0 \quad (17)$$

The MC is higher with improved seeds because of the increase in the probability of the tax. By contrast, the difference in MB can be positive, negative or null as it depends on the shape of the utility function and the difference in marginal productivity of labour between improved and local seeds, $b_I - b_L$. In other terms, improved seeds might change the decision from asking to not asking because the marginal benefit, MB, decreases.

This is one of the assumptions we will make in the empirical section: the improved seeds do not decrease the MB and therefore a decrease in the number of social interactions is an evidence of hiding.

Furthermore, as we assumed that the network size doesn't enter in the marginal benefit of asking, when n increases, the difference in the MCs increases while the difference in the MBs does not:

$$\frac{\partial \Delta MB}{\partial n} = 0 \quad (18)$$

$$\frac{\partial \Delta MC}{\partial n} = \frac{\partial \pi(n,k,f)}{\partial k \partial n} U'(c_F - d)d > 0 \quad (19)$$

Therefore, we expect hiding to increase as the size of the social network increases:

$$\frac{\partial \Delta V}{\partial n} \approx -\frac{\partial \pi(n,k,f)}{\partial k \partial n} U'(c_F - d)d < 0 \quad (20)$$

The social network, n, is only included in the cost side of the model presented in equation (1) and (2), i.e. only via an increase in the pressure to share. However, the literature on social networks has provided many evidence of positive effects of social networks (for a recent literature review, see Chuang and Schechter 2015). It is only recently that studies pointed to a 'dark side' of the social network (Di Falco and Bulte 2011). The network should therefore also be included in the benefit side of equation (1) and (2).

For instance, asking for help on the farm is a good way of nurturing the relationships with the members of the social network: during fields' work, farmers share stories and reinforce friendship bonds. The farmer hence increases his chances of benefiting from his friends' solidarity in case of a downturn, such as illness or crop damage caused by pest invasion. Friends could also provide credit for investment on the farm or gift on special occasions such as weddings or funerals. Hence, asking for help on the farm does not only provide the tangible benefit of an increase in the labour supply, it also helps fostering strong solidarity bonds which might provide informal source of insurance and credit. An additional benefit could hence be added in equation (1) and (2).

Furthermore, the size of the social network and the utility derived from the social network might be interdependent: farmers with larger network might value more social interactions. They would hence tend to interact more with other people, and hence have a larger social network. In other terms, the benefit of asking for help might increases with the size of the social network:

$$\frac{\partial MB}{\partial n} > 0 \tag{21}$$

Lastly, social networks have been shown to play an important role in technology diffusion by increasing farmers' awareness of new agricultural technologies. Farmers with a larger social network might hence have a better know-how of improved seeds cultivation. The increase on harvest brought by a helping hand, b_I, might hence increase with the size of the social network. Therefore, the difference in MB as n increases might hence be positive:

$$\frac{\partial \Delta MB}{\partial n} > 0 \tag{22}$$

Once the benefit side of the social network is taken into account, the decision to hide does not automatically increases as the size of the social network increase:

$$\frac{\partial \Delta V}{\partial n} \approx \frac{\partial \Delta MB}{\partial n} - \frac{\partial \Delta MC}{\partial n} \lesseqgtr 0 \tag{23}$$

because both $\frac{\partial \Delta MC}{\partial n}$ and $\frac{\partial \Delta MB}{\partial n}$ are positive. It depends on how ΔMB and ΔMC vary with the network size. One of the assumptions we will make, when interpreting the results, is hence that the difference in the marginal benefit of asking between farmers with improved and local seeds doesn't decrease as the social network increases.

As mentioned in section 4.2, we will also investigate other types of social interactions. These social interactions entail different degrees of visibility and different degrees of benefits. For instance, discussing the type of seeds has a high visibility but no direct tangible impact on the harvest of the experimental plot, i.e. it does not bring any b_I but it increases the difference in MC by $\frac{\partial \pi(n,k,f)}{\partial k}$. Asking for information about land and output markets or asking for help in cash or in kind do not imply a large visibility nor directly benefit the plot so it should not cause any differences in MB or MC.

As a final note, farmers could choose to substitute highly visible social interactions by less visible ones: chatting more about markets rather than about the type of seeds for instance. Therefore, hiding from peer pressure by not asking for help on the farm does not imply that the farmers renounce the benefits provided by social networks. Furthermore, refraining from asking for help on the farm might go unnoticed and is unlikely to be regarded as a violation of the social norms regulating moral economies.[11] Indeed, much more proactive and blatant hiding strategies have been observed. For instance, Baland (2011) found reports that individuals in Cameroon contracted unneeded debts in order to pretend to be poor and fend off pressure to share, while Anderson and Baland (2002) found that women in Kenyan slums joined rotating savings and credit associations in order to protect their savings from their husbands.

4.4 Econometric strategy

We start by considering social interactions which increase the risk of revealing the type of seeds received in the experiment, i.e. implying a high visibility: discussing the type of seeds received in the experiment, hiring someone for harvesting the experimental plot, or asking for help on the farm. The goal of the estimation is to test if farmers having received improved seeds seek to hide this positive shock from their kin, i.e. if they seek to escape the social pressure to share. We consider that farmers having received improved seeds hide if they decrease the number of social interactions compared to farmers with local seeds and the same network size. As detailed in section 4.3, we expect that the social pressure to share increases with number of kin. The hiding behaviour should therefore be greater for farmers with a larger social network.

In order to test these hypotheses, we estimate the following regression line:

$$D = \beta_0 + \beta_S S + \beta_N N + \beta_I I + \boldsymbol{C}'\boldsymbol{\beta} \tag{24}$$

[11] By contrast, refusing to provide help could constitute a violation of the 'social contract' (Hoff and Senn, 2006) and could lead to large social sanctions: 'To fail in kinship obligation is to be a witch…, in other words to be the opposite of a moral being: a murderer, a bestialist, a lover of death, etc.' (Bloch, 1974).

where S is a dummy variable equal to 1 if the farmer belongs to the group with improved seeds and zero otherwise (local seeds), N is the network size, $I = N * S$ the interaction effect between the improved seeds dummy and the network size. We also add a set of controls, $\boldsymbol{C}$, summarized in section 4.2. D is either the number of people with whom each farmer discussed the type of the seeds received in the experiment, or the number of people asked to help on the farm in general or on the experimental plot at harvest time. We run hence three separate set of regressions, one for each dependent variable.

Under the null hypothesis that farmers with improved seeds did not attempt to hide, the marginal effect of improved seeds should be equal to zero:

$$H_0: \hat{\beta}_S + \hat{\beta}_I N = 0 \tag{25}$$

Furthermore, under the null hypothesis that hiding does not increase with the size of the social network, then $\hat{\beta}_I$ should be not statistically different from zero. We can expect that the β_N is positive, i.e. farmers with local seeds discuss with more people the type of seeds they got if their social network is larger.

Social interactions which do not involve visibility should not differ between farmers with improved and local seeds as theses social interaction do not increase the risk of a family tax. In order to test for this, the left-hand side variable of equation (24) is replaced with either the number of people asked for help in kind and in cash, or the number of people asked about information on output and land market, or about best farming practices.

Although the seeds' allocation to farmers was randomized and is hence totally exogenous, the network variable, N, might be correlated with some unobserved heterogeneity such as a long history of asking and giving help or information, social capital, socio-demographic characteristics or farm management skills. We seek to control for this by adding a set of control variables.

Nevertheless, some important factors left in the error term might still be correlated with the social network variable. A priori, this could be a serious concern since the resulting endogeneity implies that the estimates of the interaction term are biased. Conventional econometric wisdom indicates that all estimates, (β_S, β_N, β_I) are biased and inconsistent because of *smearing* (e.g. Greene 2003). However, several authors have recently used interaction effects between exogenous and endogenous variables (Abhijit V. Banerjee et al. 2007; Abhijit V Banerjee et al. 2010; Glewwe et al. 2009). Nizalova & Murtazashvili (2014) have shown both analytically and with simulations that the OLS estimate of the interaction effect is biased but consistent. The condition is that the endogenous variable and the unobserved heterogeneity to which the former is correlated are jointly independent from the exogenous treatment, a condition fulfilled thanks the randomization of seed allocation. We assume that our sample is large enough to yield consistent estimates. We present in appendix A a step by step derivation of the main results of Nizalova and Murtazashvili (2014) in order to offer a comprehensive treatment of our identification strategy. In order to limit the effect of having many zeroes in the dependent variable, i.e. people who have not asked for help, we will use a zero-inflated Poisson model alongside standard OLS regressions. All regression results are presented with standard error, robust to clustering at the village level.

4.5 Results

We start by analysing a social interaction entailing high visibility and low benefit: discussing the type of seeds received in the experiments. The dependent variable is hence the number of people with whom each farmer discussed the type of seeds received in the experiment, hereafter number of discussions. We estimated a first Ordinary Least Squares (OLS) model where the dependent variable is log-transformed because it is highly skewed. Furthermore, a large number of households did not discuss the type of seeds at all. We therefore also estimate a zero-inflated Poisson (ZIP) model, more appropriate than an OLS when analysing count data with an excess number of zeroes[12].

We tested several specifications. In model (1), the only variables are the improved seeds dummy and the number of relative within the village (see section 4.2 for the definition). In model (2), we add an interaction term between the improved seeds dummy and the number of relatives. In model (3), we add a dummy equal to one for farmers in the north as well as a series of control variables for farmers' social capital: leadership role in the community (dummy), member of a self-help group (dummy) and member of a religious association (dummy). In model (4), we add a series of controls on farm and socio-demographic characteristics and on the farm inputs: experimental plot size (ha), walking distance to the plot (minutes), farm size (ha), oxen (dummy), labour (man day), pest damage (dummy), standardized precipitation index, female headed household (dummy), age of the household head, household size, secondary education (dummy), risk averse (dummy). The goal of this set of variable is to limit the risk of an omitted variable bias. Table 4.4 shows the results.

[12] We tested as an alternative to the ZIP model the Poisson model, but results of the Vuong test show unambiguously that the ZIP perform better whatever control variables are added.

Table 4.4: Discussing the type of seeds

Determinants of the Number of People with Whom Each Farmer Discussed the Type of Seeds Received in the Experiment

	Ordinary Least Squares Estimates: log-linear model				*Zero-Inflated Poisson Estimates*			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Log count								
Improved seeds (d)	-0.00	0.22	0.21	0.14	-0.20**	0.05	0.03	-0.03
	(0.09)	(0.13)	(0.13)	(0.13)	(0.09)	(0.21)	(0.21)	(0.19)
Relatives	0.13***	0.23***	0.22***	0.16***	0.10***	0.18**	0.17**	0.14*
	(0.02)	(0.04)	(0.04)	(0.04)	(0.04)	(0.08)	(0.09)	(0.08)
Improved seeds (d) *Relatives		-0.20***	-0.16***	-0.14***		-0.13*	-0.13	-0.11
		(0.04)	(0.04)	(0.03)		(0.08)	(0.09)	(0.07)
North (d)			X	X			X	X
Social capital controls			X	X			X	X
Inputs and socio-demographic controls				X				X
Constant	1.00***	0.87***	0.97***	0.82***	1.39***	1.23***	1.34***	0.77**
	(0.08)	(0.11)	(0.10)	(0.20)	(0.14)	(0.21)	(0.19)	(0.33)
Log odd of asking no one (logit model)								
Improved seeds (d)					-0.18	-0.92*	-1.12*	-1.19
					(0.34)	(0.51)	(0.64)	(0.76)
Relatives					-0.21	-0.62**	-0.70**	-0.73
					(0.13)	(0.30)	(0.34)	(0.46)
Improved seeds (d) *Relatives						0.65*	0.72*	0.84
						(0.36)	(0.41)	(0.54)
North (d)							X	X
Social capital controls							X	X
Inputs and socio-demographic controls								X
Constant					-1.41***	-0.96***	-1.95**	-1.64***
					(0.35)	(0.36)	(0.82)	(0.61)
Observations	298	300	299	297	313	313	313	313
Adjusted R^2	0.08	0.11	0.11	0.23				
F-test	25.4***	20.2***	46.5***	5.5***				
AIC	595	599	592	551	1910	1892	1815	1755
BIC	606	614	622	607	1932	1922	1867	1811
McFadden's Pseudo R^2					0.03	0.04	0.082	0.114
Wald chi^2					18.7***	21.2***	20.8***	58.6***

*Cluster robust standard errors at the village level in parentheses, * p < 0.1, ** p < 0.05, *** p < 0.01, (d) stands for dummy variables. The social network variable is 'Relatives': it is the average number of relatives within the village each farmer can ask for help or information. List of social capital controls: leadership role in the community (d), member of a self-help group(d), member of a religious association(d). List of inputs and socio-demographic controls: plot size (ha), walking distance to the plot (minutes), farm size (ha), oxen (d), labour (man day), pest damage (d),* *standardized precipitation index, female headed household (d), age, household size, secondary education (d), risk averse (d).*

Model (1) shows that receiving improved seeds has no significant effect on the average number of discussions. The only significant variable is the number of relatives within the village: when the number of relatives increases by one, the number of discussions increases by 13%. The addition in model (2) of the interaction term between receiving improved seeds and the number of relatives increases slightly the model fit (from 0.08 to 0.11)[13]. The interaction term is highly significant and is negative as expected: while farmers with local seeds discuss more if they have numerous kin, it is not the case for farmers with improved seeds. The addition of controls in model (3) and (4) does not affect the results: farmers with local seeds discuss more if they count more relatives, farmers with improved seeds do not. The ZIP models corroborate these findings. The number of relatives decreases the odd of not discussing (logit part of the ZIP model), but less so for farmers with improved seeds. The number of relatives increases the number of discussions (Poisson part of the ZIP model), but less so for farmers with improved seeds.

Figure 4.2: Discussing the type of seeds

[13] However, both the Akaike and Bayesian Information Criterion (AIC and BIC) increases. Nevertheless, as the main interest of the current analysis is in the sign of the interaction term coefficient, we choose to keep it in the model. Furthermore, the AIC and the BIC decreases from models (5) to (6) in the ZIP model estimation.

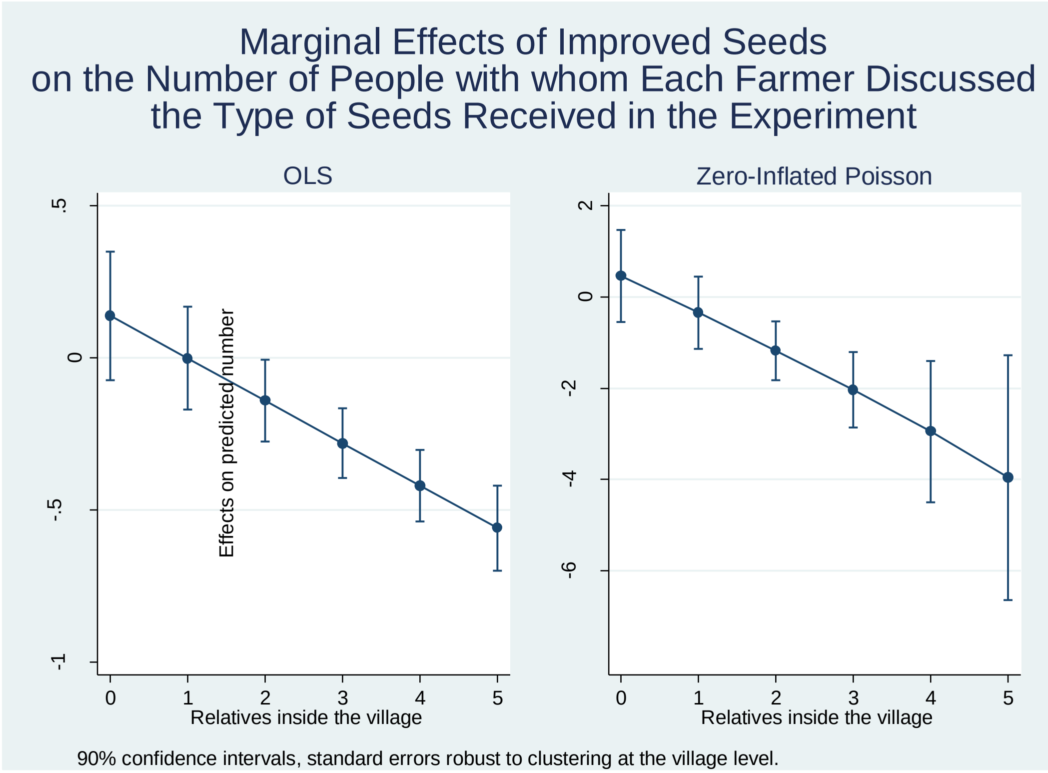


Figure 4.2 shows the marginal effects of receiving the improved seeds on the number of discussions estimated according to the OLS model (4) and the ZIP model (8). We define hiding as a decrease in the number of discussions caused by the improved seeds.

Note that there is no evidence of hiding among farmers counting only one or no relatives within the village. This is expected, as when the social pressure to share is low, there is no incentive to hide. By contrast, if farmers count two relatives, those with improved seeds discussed 12% less than those with local seeds. Lastly, if they count more than 5 relatives in the village, i.e. when the pressure to share is very high, they discuss 50% less when they receive improved seeds (OLS results). Similarly, the ZIP estimates shows that hiding becomes statistically significant when farmers count more than 1.5 relatives within the village, i.e. it is significant for 40% of the sample. The sample average effect is large: receiving improved seeds decreases by 24% the number of discussions (ZIP model, (8)). Lastly, the graphs show that the incentive to hide builds up at the number of relatives increases. This is coherent with the hypothesis that larger is the number of relatives, larger is the pressure to share and hence larger is the incentive to hide.

We now investigate another social interaction increasing the risk of disclosing the type of seeds received in the experiment: asking for help on the farm. Given the large number of farmers who did not ask anyone for help on the farm (43%), we estimated only a ZIP model. The set of controls is the same, except that here we also control for having been asked for help on the farm of others in model (4) and (5). Table 4.5 shows the results.

Table 4.5: Asking for help on the farm

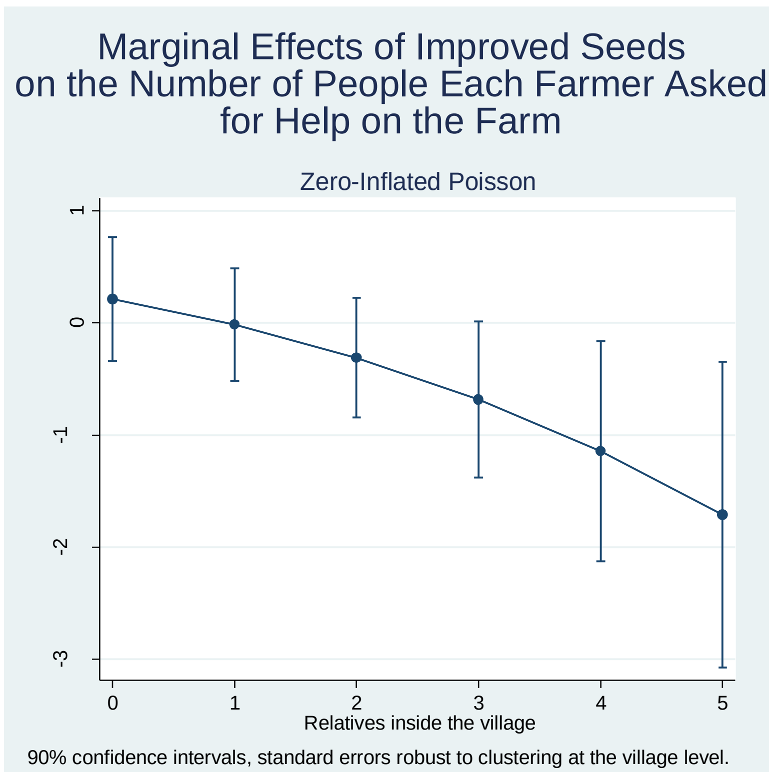
Zero-Inflated Poisson Estimates: Number of People Each Farmer Asked to Come for Help on the Farm					
	(1)	(2)	(3)	(4)	(5)
Log count (Poisson)					
Improved seeds (d)	-0.15	0.01	0.01	0.05	0.14
	(0.12)	(0.24)	(0.24)	(0.20)	(0.22)
Relatives	0.06	0.11***	0.11***	0.14***	0.14***
	(0.05)	(0.04)	(0.04)	(0.03)	(0.03)
Improved seeds (d) *Relatives		-0.09	-0.09	-0.12**	-0.14**
		(0.09)	(0.09)	(0.06)	(0.06)
North (d)			X	X	X
Social capital controls				X	X
Inputs and socio-demographic controls					X
Constant	1.13***	1.04***	0.97***	0.81***	1.30***
	(0.12)	(0.16)	(0.27)	(0.21)	(0.46)
Log odd of asking no one (logit)					
Improved seeds (d)	-0.02	-0.07	-0.09	0.21	0.03
	(0.21)	(0.44)	(0.44)	(0.34)	(0.38)
Relatives	-0.23***	-0.25	-0.27	-0.28	-0.43
	(0.07)	(0.18)	(0.17)	(0.20)	(0.30)
Improved seeds (d) *Relatives		0.05	0.08	-0.08	0.03
		(0.26)	(0.28)	(0.23)	(0.28)
North (d)			X	X	X
Social capital controls				X	X
Inputs and socio-demographic controls					X
Constant	-0.05	-0.03	0.38	1.68***	2.28***
	(0.27)	(0.41)	(0.33)	(0.49)	(0.84)
Observations	311	311	311	311	311
McFadden's Pseudo R^2	0.0	0.02	0.02	0.174	0.194
Wald chi^2	2.22	12.28***	13.10***	28.98***	47.84***
AIC	1256	1257	1252	1071	1045
BIC	1278	1287	1290	1127	1101

*Cluster robust standard errors at the village level in parentheses, * p < 0.1, ** p < 0.05, *** p < 0.01, (d) stands for dummy variables. The social network variable is 'Relatives': it is the average number of relatives inside the village each farmer can ask for help or information. List of social capital controls: leadership role in the community (d), member of a self-help group(d), member of a religious association(d), being asked for help on the farm of other (d). List of inputs and socio-demographic controls: plot size (ha), walking distance to the plot (minutes), farm size (ha), oxen (d), labour (man day), pest damage (d), standardized precipitation index, female headed household (d), age, household size, secondary education (d), risk averse (d).*

The pattern is similar to the one observed in the case of discussing the seeds. The logit model shows that for farmers with local seeds the odds of not asking anyone decreases with the number of relatives, i.e.

farmers with a larger number of relatives have a higher propensity to ask than those with smaller number of relative. It is also true of farmers with improved seeds, but the increase in the propensity of asking as the network size increases is smaller. The Poisson part of the models shows a stronger pattern. The number of relatives increases the predicted number of people asked to help on the farm for farmers having received the local seeds while the effect of the number of relatives is statistically not different from zero for farmers with improved seeds. We summarize the marginal effects of receiving improved seeds in Figure 4.3 .

Figure 4.3: Asking for help on the farm



There is less evidence of hiding than in the case of discussing the type of seeds. This is expected as not discussing the type of seeds is a relatively cost-free hiding strategy while not asking for help on the farm implies either that the household has to increase his own labour provision or that some of the farm tasks wo not be performed as well as with the help from others (e.g. less care in planting and weeding). Nevertheless, we do find that hiding increases as the number of kin increases. It becomes statistically significant for farmers with three or more relatives within the village, i.e. 15% of the sample. The hiding behaviour appears hence to take place only among farmers exposed to the largest pressure to redistribute.

Farmers taking part to the experiment, as across many places in Sub-Saharan Africa, tends to have several plots which can be far away from each other. Therefore, some farmers may ask for help on plots other than the plot where they planted the seeds of the experiment. We turn now to the decision to ask for help on the experimental plot.

We start by investigating the decision to ask for help at harvest time, i.e. when there is the highest visibility. The dependent variable is a dummy variable equal to one if the farmer asked for help on the farm and zero otherwise. We used the same set of explicative variables than presented above. Based on the comparison of the log-likelihood, a logit model is preferred to a probit model. Table 4.6 shows the results. Detailed results are left in the appendix.

Table 4.6: Asking for help to harvest the experimental plot

Logit Estimates: Asking for Help on the Experimental at Harvest times					
	(1)	(2)	(3)	(4)	(5)
Improved seeds (d)	-0.01	0.78	0.77	0.66	0.42
	(0.29)	(0.54)	(0.55)	(0.55)	(0.35)
Relatives	0.28***	0.53***	0.53***	0.53***	0.39***
	(0.10)	(0.16)	(0.15)	(0.16)	(0.12)
Improved seeds (d) *Relatives		-0.38**	-0.38**	-0.37**	-0.29**
		(0.16)	(0.16)	(0.15)	(0.12)
North (d)			X	X	X
Harvest (10 kg)				X	X
Controls				X	X
Constant	-2.67***	-3.21***	-3.43***	-3.69***	-5.94***
	(0.39)	(0.57)	(0.66)	(0.74)	(1.64)
Observations	313	313	313	313	313
McFadden's Pseudo R^2	0.04	0.05	0.06	0.06	0.17
Wald chi^2	7.61**	14.3***	21.6***	29.3***	41.20***
AIC	205	204	205	206	201
BIC	216	219	224	228	257

Cluster robust standard errors at the village level in parentheses, * p < 0.10, ** p < 0.05, *** p < 0.01, (d) stands for dummy variables. The social network variable is 'Relatives': it is the average number of relatives inside the village each farmer can ask for help or information. List of controls: plot size (ha), farm size (ha), oxen (d), pest damage (d), standardized precipitation index, female headed household (d), age, household size, secondary education (d), risk averse (d), leadership role in the community (d), member of a self-help group(d) member of a religious association(d) walking distance to the plot (minutes).

As expected, an increase in the number of relatives increases the odds of asking for help at harvest time. Indeed, a larger kin network increases the opportunity to enter labour sharing agreement where labour is pooled and the harvest is done in common, one field after the other. However, there is a clear distinction between farmers having received the improved and the local seeds. Among the formers ones, farmers with a large network will ask less on average than those with a small network.

We also tested if we could find a similar effect on the decision to ask for help for pre-harvest work on the experimental plot (e.g. planting, weeding or slashing). The marginal productivity of labour on improved seeds fields might higher than on local seeds plots. Therefore, the marginal benefit of the helping hand might outweigh the marginal cost of increased visibility. Furthermore, in this earlier phase of the crop growing cycle, harvest is more difficult to ascertain and adverse weather or a pest invasion can still deteriorate the harvest. Therefore, asking for help in pre-harvest entails less visibility than at harvest time. We should therefore observe a lower effect of improved seeds on the decision to ask for help. Detailed results are left in the appendix in Table 4.11. The results are summarized in Figure 4.4 (left hand-side) alongside the marginal effect of improved seeds on asking for help at harvest time (right-hand side).

Figure 4.4: Asking for help on the experimental plot

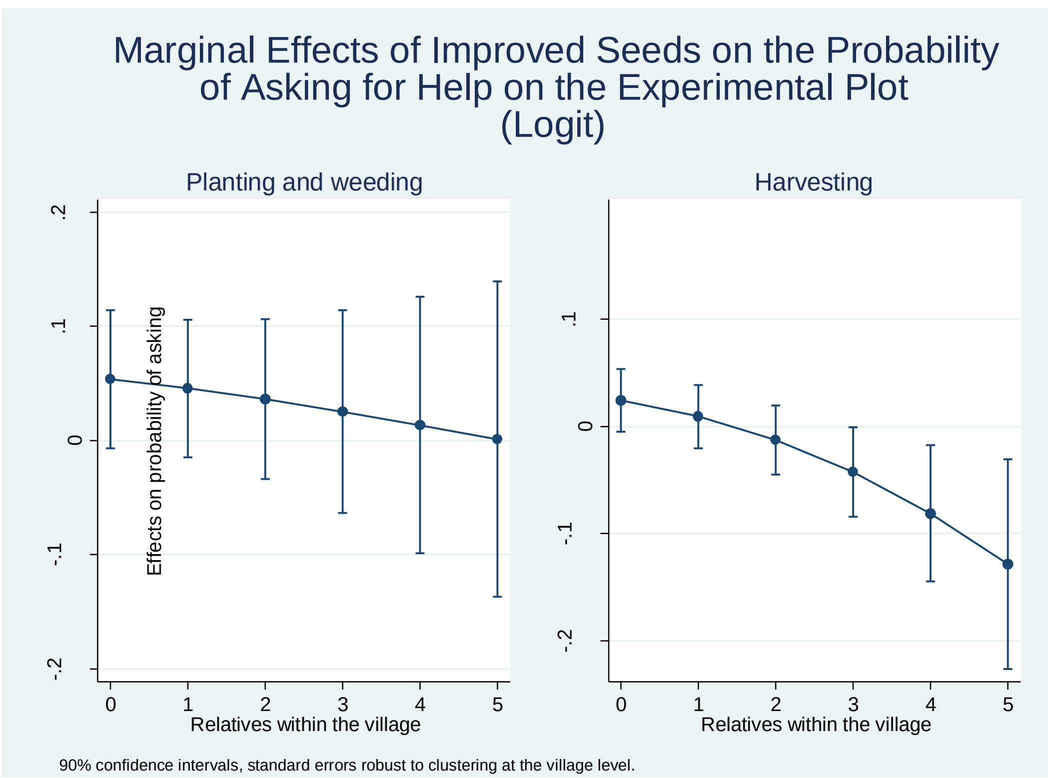


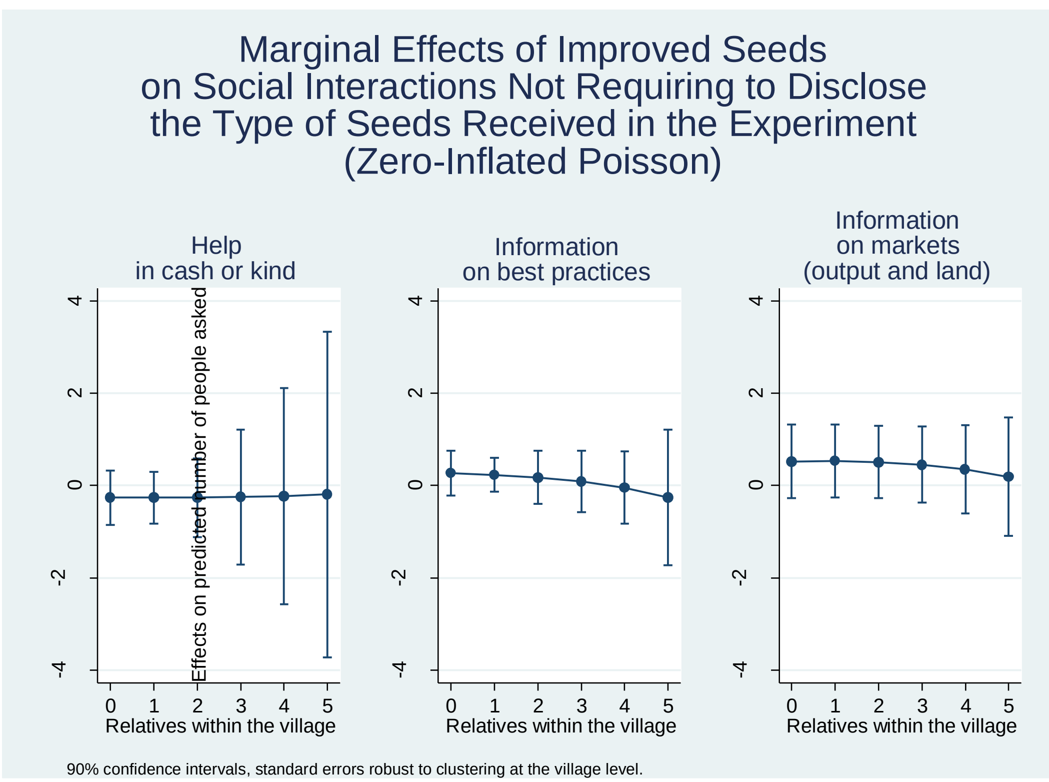
Figure 4.4 shows that while improved seeds have no impact on the decision to ask for pre-harvest work however large the number of relatives within the village, this is not the case anymore when it comes to help at harvest time. In the latter case, there is clearly a reduction in the probability of asking as soon as

the number of relative is large, i.e. there is evidence of a hiding behaviour. The fact that farmers with improved seeds and a large social network do not hesitate to ask in pre-harvest times while they abstain from it at harvest time suggests that decreasing the visibility of their harvest is important for them. Hiding becomes significant for farmers counting at least three relatives in the village, i.e. 15% of the sample.

We now test if we can find hiding behaviours in social interactions which are not expected to increase the risk of a family tax: (1) the number of people asked for help in cash or in kind; (2) the number of people asked about information on best farming practices; (3) the number of people asked about information on output or land market. Naturally, it could be argued that these social interactions involve also some sort of visibility. However, it is much more limited than discussing the seeds received in the experiments and asking people for help on the farm.

The same set of explicative variables is used to analyse the determinants of these social interactions. We also used a ZIP model in order to model the excess zeroes. Detailed results are present in appendix in Table 4.11. Figure 4.5 summarizes the results.

Figure 4.5: Social interactions not related to the experimental plot



We do not find any difference between farmers having received improved and local seeds in terms of the probability of engaging in social interaction not related to the experimental plot. In other terms, there is no sign of hiding behaviour however large the kin network is. Allocating farmers to the improved seeds group did therefore distort social interactions only in the case where there was a risk of an increase in the family tax, i.e. social interactions revealing the positive consumption shock brought about by improved seeds distributed in the experiment.

4.6 Conclusion

This chapter tests the hypothesis that individuals try to escape forced solidarity when facing favourable conditions. We randomly assigned a positive shock by providing farmers with improved maize seeds that are more productive than the traditional varieties. Our results, robust to various specifications, show that farmers who are assigned more productive seeds decrease the number of their social interactions if they face a social pressure to share, i.e. if they count many relatives within the village. This suggests that farmers attempt to reduce the burden of decreasing their future consumption in order with other members of the social network, which is consistent with the idea that traditional sharing norms may invite evasive behaviour.

The results of the present chapter provide another set of evidence of the existence of evasive responses to the social pressure to share. Di Falco and Bulte (2011) find that forced sharing norms in social networks diverted investment away from sharable liquid assets and could have a negative effect on growth while Jakiela and Ozier (2016) find that sharing norms distorted incentives ‘towards less visible, but potentially less profitable, investments, and may consequently slow economic growth’.

This dark side of social capital echoes the concept of the Laffer curve: increasing tax level increases government revenues up to a point where work is discouraged, tax evasion kicks in and government revenues drop. A similar phenomenon could be at play in the case of village economies where solidarity networks play a central role in income redistribution and forced solidarity is akin to a tax. Some farmers might prefer to forsake an increase in output brought about by asking for help in order to reduce the likelihood of the tax, particularly if the likelihood of a tax is high because their network is large and their wealth status has improved. Indeed, why ask for help if the resulting increase in output will be eaten away by friends and family? It might be better to keep a low profile, enjoy the bumper crop harvest brought about by the improved seeds and relinquish an even greater harvest as the latter would be taxed by members of their social network. Hence, as in the Laffer curve hypothesis, where higher tax brings about smaller revenues, the collision of self-interest and high solidarity might imply that less food is produced and available for sharing at the village level. This hypothesis could be investigated in future work.

Indeed, it is yet not clear if the hiding behaviour identified in the present study has any economic consequences. We do indeed find that a farmer with improved seeds asks for less help on the farm if their kin network is large, but the hiding behaviour only becomes significant for the 15% of farmers with the largest number of relative within the village. Furthermore, as no hiding was found in terms of the pre-harvest labour, i.e. when labour is expected to have the highest effect on harvest, it is yet not clear if this hiding strategies implied a decrease in output. In our experiment, farmers may have been able to hide at no cost.

4.7 References

Alger, Ingela and Weibull, Jörgen W (2010), 'Kinship, incentives, and evolution', *The American Economic Review*, 1725-58.

Anderson, Siwan and Baland, Jean-Marie (2002), 'The economics of roscas and intrahousehold resource allocation', *Quarterly Journal of Economics*, 963-95.

Baland, Jean-Marie, Guirkinger, Catherine, and Mali, Charlotte (2011), 'Pretending to be poor: Borrowing to escape forced solidarity in Cameroon', *Economic Development and Cultural Change,* 60 (1), 1-16.

Banerjee, Abhijit V, et al. (2010), 'The miracle of microfinance? Evidence from a randomized evaluation', *BREAD Working Paper,* 278.

Banerjee, Abhijit V., et al. (2007), 'Remedying Education: Evidence from Two Randomized Experiments in India', *The Quarterly Journal of Economics,* 122 (3), 1235-64.

Barr, Abigail and Stein, Mattea (2008), 'Status and egalitarianism in traditional communities: An analysis of funeral attendance in six Zimbabwean villages', *CSAE WPS 26*.

Chuang, Yating and Schechter, Laura (2015), 'Social Networks in Developing Countries', *Annual Review of Resource Economics,* 7.

Coate, Stephen and Ravallion, Martin (1993), 'Reciprocity without commitment: Characterization and performance of informal insurance arrangements', *Journal of development Economics,* 40 (1), 1-24.

Dercon, Stefan (2002), 'Income risk, coping strategies, and safety nets', *The World Bank Research Observer,* 17 (2), 141-66.

Di Falco, Salvatore and Bulte, Erwin (2011), 'A dark side of social capital? Kinship, consumption, and savings', *Journal of Development Studies,* 47 (8), 1128-51.

Fafchamps, Marcel (1992), 'Solidarity networks in preindustrial societies: Rational peasants with a moral economy', *Economic development and cultural change*, 147-74.

Fafchamps, Marcel and Lund, Susan (2003), 'Risk-sharing networks in rural Philippines', *Journal of development Economics,* 71 (2), 261-87.

Fafchamps, Marcel and Gubert, Flore (2007), 'The formation of risk sharing networks', *Journal of Development Economics,* 83 (2), 326-50.

Foster, Andrew D and Rosenzweig, Mark R (2001), 'Imperfect commitment, altruism, and the family: Evidence from transfer behavior in low-income rural areas', *Review of Economics and Statistics,* 83 (3), 389-407.

Glewwe, Paul, Kremer, Michael, and Moulin, Sylvie (2009), 'Many children left behind? Textbooks and test scores in Kenya', *American Economic Journal: Applied Economics,* 1 (1), 112-35.

Greene, William H (2003), *Econometric analysis* (Pearson Education India).

Hoff, Karla and Sen, Arijit (2006), 'The Kin system as a poverty trap', in Cayley Bowles, Steven N. Durlauf, and Karla Hoff (eds.), *Poverty Traps* (Princeton: Princeton University Press).

Jakiela, Pamela and Ozier, Owen (2016), 'Does Africa Need a Rotten Kin Theorem? Experimental Evidence from Village Economies', *The Review of Economic Studies,* 83 (1), 231-68.

Kazianga, Harounan and Udry, Christopher (2006), 'Consumption smoothing? Livestock, insurance and drought in rural Burkina Faso', *Journal of Development Economics,* 79 (2), 413-46.

Leider, Stephen, et al. (2009), 'Directed altruism and enforced reciprocity in social networks', *The Quarterly Journal of Economics,* 124 (4), 1815-51.

Ligon, Ethan (2011), 'Cooperation, Reciprocity and Altruism in Developing Countries: Discussion', *American Journal of Agricultural Economics*, aar104.

Ligon, Ethan and Schechter, Laura (2012), 'Motives for sharing in social networks', *Journal of Development Economics,* 99 (1), 13-26.

Ligon, Ethan, Thomas, Jonathan P, and Worrall, Tim (2002), 'Informal insurance arrangements with limited commitment: Theory and evidence from village economies', *The Review of Economic Studies,* 69 (1), 209-44.

Nizalova, Olena Y and Murtazashvili, Irina (2014), 'Exogenous treatment and endogenous factors: Vanishing of omitted variable bias on the interaction term', *Journal of Econometric Methods*.

Novella, Nicholas S and Thiaw, Wassila M (2013), 'African rainfall climatology version 2 for famine early warning systems', *Journal of Applied Meteorology and Climatology,* 52 (3), 588-606.

Platteau, Jean-Philippe (1991), 'Traditional Systems of Social Security and Hunger Insurance: Past Achievements and Modern Challenges', in Ehtisham Ahmad, et al. (eds.), *Social Security in Developing Countries* (Oxford: Clarendon Press).

--- (2000), *Institutions, social norms, and economic development* (1: Psychology Press).

Rosenzweig, Mark R (1988), 'Risk, implicit contracts and the family in rural areas of low-income countries', *The Economic Journal*, 1148-70.

Scott, James C (1976), 'The moral economy of the peasant: Subsistence and rebellion in Southeast Asia', *Yale UP (New Haven)*.

Suri, Tavneet (2011), 'Selection and comparative advantage in technology adoption', *Econometrica,* 79 (1), 159-209.

Townsend, Robert M (1994), 'Risk and insurance in village India', *Econometrica: Journal of the Econometric Society*, 539-91.

Udry, Christopher (1994), 'Risk and insurance in a rural credit market: An empirical investigation in northern Nigeria', *The Review of Economic Studies,* 61 (3), 495-526.

4.8 Appendix A: Interaction effect, biased but consistent

We reproduce below a step by step derivation of the results of Nizalova and Murtazashvili (2014) in order to offer a comprehensive treatment of the identification strategy of the present paper. We want to estimate the following model:

$$Y = \beta_0 + \beta_I I + \beta_N X_N + \beta_S X_S + \varepsilon^* \tag{26}$$

where $I = X_S X_N$, $\varepsilon^* = \varepsilon + \beta_H H$ and X_S is perfectly exogenous. The parameter of interest is β_I while the issue is the correlation between H and X_N. The omitted variable bias is given by (e.g. Greene 2003):

$$plim(\hat{\beta}) = \beta + \gamma \boldsymbol{Q} \tag{27}$$

where $\boldsymbol{Q} = plim[(X'X)^{-1}(X'X)\varepsilon^*] = plim[(X'X)^{-1}(X'H)]$.

Let's focus on $\boldsymbol{Q}$. We start by re-expressing $X'X$:

$$X'X = \begin{pmatrix} \sigma_{I1}^2 & \sigma_{IN} & \sigma_{IS} \\ \sigma_{NI} & \sigma_N^2 & \sigma_{NS} \\ \sigma_{SI} & \sigma_{SN} & \sigma_S^2 \end{pmatrix} = \Lambda R \Lambda \tag{28}$$

where $\Lambda = \begin{pmatrix} \sigma_{I1}^2 & 0 & 0 \\ 0 & \sigma_N^2 & 0 \\ 0 & 0 & \sigma_S^2 \end{pmatrix}$, $R = \begin{pmatrix} 1 & r_{IN} & r_{IS} \\ r_{NI} & 1 & r_{NS} \\ r_{SI} & r_{SN} & 1 \end{pmatrix}$ and $r_{i,j}$ is the correlation between variable i and variable j, $r_{i,j} = \frac{\sigma_{i,j}}{\sigma_i \sigma_j}$.

The second pair of matrices can be expressed as:

$$X'H = \sigma_H \Lambda \omega \tag{29}$$

where $\omega' = [r_{IH}, r_{NH}, r_{SH}]$. Hence, $\boldsymbol{Q} = plim[\sigma_H (\Lambda R \Lambda)^{-1} \Lambda \omega] = plim[\sigma_H \Lambda^{-1} \mathrm{R}^{-1} \omega]$. We note that:

$$\mathrm{R}^{-1} = \frac{1}{\det(R)} \begin{pmatrix} 1 - r_{NS}^2 & r_{IS} r_{NS} - r_{IN} & r_{IN} r_{NS} - r_{IS} \\ r_{IS} r_{NS} - r_{IN} & 1 - r_{IS}^2 & r_{IS} r_{IN} - r_{NS} \\ r_{IN} r_{NS} - r_{IS} & r_{IS} r_{IN} - r_{NS} & 1 - r_{IN}^2 \end{pmatrix} \tag{30}$$

where $\det(R) = 1 - r_{IN}^2 - r_{IS}^2 - r_{SN}^2 + 2 r_{IN} r_{IS} r_{NS}$. Hence, we can express $\boldsymbol{Q}$ as a vector of size 3:

$$\boldsymbol{Q} = \begin{pmatrix} \frac{\sigma_H}{\sigma_I} \frac{r_{IH}(1 - r_{NS})^2 + r_{NH}(r_{IS} r_{NS} - r_{IN}) + r_{SH}(r_{IN} r_{NS} - r_{IN})}{1 - r_{IN}^2 - r_{IS}^2 - r_{SN}^2 + 2 r_{IN} r_{IS} r_{NS}} \\ \dots \\ \dots \end{pmatrix} \tag{31}$$

where we omitted the second and third row as we are interested here in β_I:

$$plim(\hat{\beta}_I) = \beta_I + \gamma_r \boldsymbol{Q_I} \tag{32}$$

where $\boldsymbol{Q_I}$ is the first row of $\boldsymbol{Q}$ and γ_r is the probability limit of the coefficient of I in an auxiliary regression of the omitted variable H on the set of covariates (I, N, S).

Assuming that X_N and H are jointly independent from the randomly assigned X_S, $r_{NS} = r_{SH} = 0$, we have:

$$\boldsymbol{Q_I} = \frac{\sigma_H}{\sigma_I} \frac{r_{IH} - r_{NH} r_{IN}}{1 - r_{IN}^2 - r_{IS}^2} \tag{33}$$

The joint independence of (X_N, H) from X_S implies that X_S conditional on X_N is independent from H. Hence, we have:

$$r_{IN} = \frac{cov(I, X_N\ )}{\sigma_I \sigma_N} \tag{34}$$

$$= \frac{E(NSN) - E(NS)E(N)}{\sigma_I \sigma_N} \tag{35}$$

$$= \frac{E(S)[E(N^2) - E(N)^2]}{\sigma_I \sigma_N} \tag{36}$$

$$= \frac{E(S)var(N)}{\sigma_I \sigma_N} \tag{37}$$

$$= \frac{\sigma_N E(S)}{\sigma_I} \tag{38}$$

From which it follows that $r_{IH} = r_{NH} r_{IH}$ and therefore $\boldsymbol{Q} = 0$ and $plim\left(\hat{\beta}_I\right) = \beta_I$.

4.9 Appendix B: Further results

Table 4.7: Discussing the seeds, details of the OLS results

Ordinary Least Squares Estimates: Number of People with whom Each Farmer Discussed the Type of Seeds Received in the Experiment					
	(1)	(2)	(3)	(4)	(5)
Improved seeds (d)	-0.00	0.22	0.22	0.21	0.14
	(0.09)	(0.13)	(0.14)	(0.13)	(0.13)
Relatives inside the village	0.13***	0.23***	0.24***	0.22***	0.16***
	(0.02)	(0.04)	(0.04)	(0.04)	(0.04)
Improved seeds (d)*Relatives inside the village		-0.20***	-0.19***	-0.16***	-0.14***
		(0.04)	(0.05)	(0.04)	(0.03)
North			-0.16*	-0.12	-0.10
			(0.07)	(0.09)	(0.11)
Leadership role in the community (d)				0.14*	0.12*
				(0.07)	(0.06)
Member of a self-help group(d)				0.12	0.08
				(0.08)	(0.09)
Member of a religious association(d)				-0.15**	-0.17**
				(0.07)	(0.07)
Plot size (ha)					2.77***
					(0.51)
Walking distance to the plot (minutes)					-0.00
					(0.00)
Farm size (ha)					0.12***
					(0.02)
Oxen (d)					0.00
					(0.03)
Labour (man day)					0.01**
					(0.01)
Pest damage (d)					0.19**
					(0.07)
Standardized Precipitation Index					0.06
					(0.07)
Female headed household (d)					-0.24*
					(0.13)
Age					-0.01**
					(0.00)
Household size					-0.02
					(0.02)
Secondary education (d)					0.12**
					(0.05)
Risk averse (d)					0.15*
					(0.07)
Constant	1.00***	0.87***	0.96***	0.97***	0.82***
	(0.08)	(0.11)	(0.13)	(0.10)	(0.20)
Observations	298	300	298	299	297
Adjusted R2	0.08	0.11	0.12	0.11	0.23
AIC	595	599	592	551	595
BIC	606	614	622	607	606

Cluster robust standard errors at the village level in parentheses, * p < 0.10, ** p < 0.05, *** p < 0.01, (d) stands for dummy variables. The social network variable is 'Relatives': it is the average number of relatives inside the *village each farmer can ask for help or information.*

Table 4.8: Discussing the seeds, details of the ZIP results

Zero-Inflated Poisson Estimates: Number of People with whom Each Farmer Discussed the Type of Seeds Received in the Experiment					
	(1)	(2)	(3)	(4)	(5)
Log count (Poisson model)					
Improved seeds (d)	-0.20**	0.05	0.06	0.03	-0.03
	(0.09)	(0.21)	(0.21)	(0.21)	(0.19)
Relatives inside the village	0.10***	0.18**	0.18**	0.17**	0.14*
	(0.04)	(0.08)	(0.08)	(0.09)	(0.08)
Improved seeds (d)*Relatives inside the village		-0.13*	-0.13	-0.13	-0.11
		(0.08)	(0.08)	(0.09)	(0.07)
North			-0.41**	-0.35**	-0.50*
			(0.17)	(0.17)	(0.30)
Leadership role in the community (d)				0.28**	0.28*
				(0.13)	(0.15)
Member of a self-help group(d)				0.01	-0.03
				(0.13)	(0.14)
Member of a religious association(d)				-0.02	-0.09
				(0.16)	(0.17)
Plot size (ha)					2.32***
					(0.77)
Farm size (ha)					-0.00
					(0.06)
Oxen (d)					-0.03
					(0.09)
Labour (man day)					0.02**
					(0.01)
Pest damage (d)					0.11
					(0.14)
Standardized Precipitation Index					0.17
					(0.22)
Female headed household (d)					-0.19
					(0.18)
Age					0.00
					(0.01)
Household size					0.03
					(0.03)
Secondary education (d)					0.04
					(0.14)
Risk averse (d)					0.17
					(0.16)
Walking distance to the plot (minutes)					0.00
					(0.00)
Constant	1.39***	1.23***	1.23***	1.34***	0.77**
	(0.14)	(0.21)	(0.21)	(0.19)	(0.33)
Log odd of asking no one (logit model)					
Improved seeds (d)	-0.18	-0.92*	-0.91	-1.12*	-1.19
	(0.34)	(0.51)	(0.58)	(0.64)	(0.76)
Relatives inside the village	-0.21	-0.62**	-0.62*	-0.70**	-0.73
	(0.13)	(0.30)	(0.32)	(0.34)	(0.46)
Improved seeds (d)*Relatives inside the village		0.65*	0.63	0.72*	0.84
		(0.36)	(0.39)	(0.41)	(0.54)
North			-0.39	-0.03	-0.86*

	(1)	(2)	(3)	(4)	(5)
			(0.32)	(0.28)	(0.46)
Leadership role in the community (d)				-0.14	0.25
				(0.25)	(0.23)
Member of a self-help group(d)				-0.56	
				(0.38)	
Member of a religious association(d)				1.62^{**}	
				(0.70)	
Farm size (ha)					-0.54^{*}
					(0.28)
Age					0.03^{*}
					(0.02)
Household size					0.16
					(0.11)
Female headed household (d)					0.93
					(0.69)
Secondary education (d)					-0.81
					(0.60)
Risk averse (d)					-0.22
					(0.68)
Constant	-1.41^{***}	-0.96^{***}	-0.96^{***}	-1.95^{**}	-1.64^{***}
	(0.35)	(0.36)	(0.36)	(0.82)	(0.61)
Observations	313	313	313	313	313
McFadden's Pseudo R^2	0.03	0.04	0.06	0.082	0.114
Wald chi^2	18.68^{***}	21.15^{***}	25.22^{***}	20.84^{***}	58.55^{***}
AIC	1910	1892	1852	1815	1755
BIC	1932	1922	1889	1867	1811

Cluster robust standard errors at the village level in parentheses, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$, (d) stands for dummy variables. The social network variable is 'Relatives': it is the average number of relatives inside the village each farmer can ask for help or information.

Table 4.9: Asking for help on the farm, details of the ZIP model

Zero-Inflated Poisson Estimates: Number of People Each Farmer Asked for Help on the Farm					
	(1)	(2)	(3)	(4)	(5)
farm_ask_no					
Improved seeds (d)	-0.15	0.01	0.01	0.05	0.14
	(0.12)	(0.24)	(0.24)	(0.20)	(0.22)
Relatives inside the village	0.06	0.11***	0.11***	0.14***	0.14***
	(0.05)	(0.04)	(0.04)	(0.03)	(0.03)
Improved seeds (d)*Relatives inside the village		-0.09	-0.09	-0.12**	-0.14**
		(0.09)	(0.09)	(0.06)	(0.06)
North			0.09	0.02	0.14
			(0.25)	(0.20)	(0.21)
Leadership role in the community (d)				0.11	0.12
				(0.15)	(0.16)
Member of a self-help group(d)				0.34***	0.37***
				(0.10)	(0.08)
Member of a religious association(d)				-0.36**	-0.45***
				(0.15)	(0.10)
Plot size (ha)					-5.00**
					(2.26)
Farm size (ha)					0.06
					(0.06)
Oxen (d)					0.08
					(0.08)
Labour (man day)					-0.00
					(0.01)
Pest damage (d)					0.14
					(0.15)
Standardized Precipitation Index					-0.12
					(0.29)
Female headed household (d)					0.02
					(0.28)
Age					0.00
					(0.00)
Household size					-0.05
					(0.05)
Secondary education (d)					-0.06
					(0.24)
Risk averse (d)					-0.02
					(0.21)
Walking distance to the plot (minutes)					-0.00
					(0.00)
Constant	1.13***	1.04***	0.97***	0.81***	1.30***
	(0.12)	(0.16)	(0.27)	(0.21)	(0.46)
inflate					
Improved seeds (d)	-0.02	-0.07	-0.09	0.21	0.03
	(0.21)	(0.44)	(0.44)	(0.34)	(0.38)
Relatives inside the village	-0.23***	-0.25	-0.27	-0.28	-0.43
	(0.07)	(0.18)	(0.17)	(0.20)	(0.30)
Improved seeds (d)*Relatives inside the village		0.05	0.08	-0.08	0.03
		(0.26)	(0.28)	(0.23)	(0.28)
North			-0.67**	-0.67	-0.60
			(0.27)	(0.44)	(0.52)

Leadership role in the community (d)				1.06**	0.00
				(0.52)	(.)
Member of a self-help group(d)				0.50	0.36
				(0.50)	(0.49)
Member of a religious association(d)				-0.51*	-0.56*
				(0.28)	(0.29)
Farm size (ha)					-0.11
					(0.18)
Age					-0.01
					(0.01)
Household size					0.06
					(0.09)
Female headed household (d)					-0.48
					(0.45)
Secondary education (d)					-0.45
					(0.34)
Risk averse (d)					0.98**
					(0.43)
Constant	-0.05	-0.03	0.38	1.68***	2.28***
	(0.27)	(0.41)	(0.33)	(0.49)	(0.84)
Observations	311	311	311	311	311
McFadden's Pseudo R^2	0.0	0.02	0.02	0.174	0.194
Wald chi^2	2.22	12.28***	13.10***	28.98***	47.84***
AIC	1256	1257	1252	1071	1045
BIC	1278	1287	1290	1127	1101

*Cluster robust standard errors at the village level in parentheses, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$, (d) stands for dummy variables. The social network variable is 'Relatives': it is the average number of relatives inside the village each farmer can ask for help or information.*

Table 4.10: Asking for help on the experimental plot a harvest times, details of the Logit model

Logit Estimates: Asking for Help on the Experimental Plot a Harvest times					
	(1)	(2)	(3)	(4)	(5)
Improved seeds (d)	-0.01	0.78	0.77	0.66	0.42
	(0.29)	(0.54)	(0.55)	(0.55)	(0.35)
Relatives inside the village	0.28***	0.53***	0.53***	0.53***	0.39***
	(0.10)	(0.16)	(0.15)	(0.16)	(0.12)
Improved seeds (d)*Relatives inside the village		-0.38**	-0.38**	-0.37**	-0.29**
		(0.16)	(0.16)	(0.15)	(0.12)
North			0.36	0.52	1.64**
			(0.48)	(0.55)	(0.81)
Harvest (10 kg)				0.00	0.00*
				(0.00)	(0.00)
Plot size (ha)					1.51
					(2.25)
Farm size (ha)					-0.18
					(0.25)
Oxen (d)					-0.49
					(0.64)
Pest damage (d)					0.57
					(0.51)
Standardized Precipitation Index					-0.17
					(0.50)
Female headed household (d)					-0.75
					(1.00)
Age					0.00
					(0.02)
Household size					-0.05
					(0.15)
Secondary education (d)					0.30
					(0.43)
Risk averse (d)					0.32
					(0.39)
Leadership role in the community (d)					0.78***
					(0.25)
Member of a self-help group(d)					0.16
					(0.40)
Member of a religious association(d)					1.74***
					(0.48)
Walking distance to the plot (minutes)					0.02***
					(0.01)
Constant	-2.67***	-3.21***	-3.43***	-3.69***	-5.94***
	(0.39)	(0.57)	(0.66)	(0.74)	(1.64)
Observations	313	313	313	313	313
Pseudo R^2	0.04	0.05	0.06	0.06	0.17
Wald chi^2	7.61**	14.3***	21.6***	29.3***	41.20***
AIC	205	204	205	206	201
BIC	216	219	224	228	257

Cluster robust standard errors at the village level in parentheses, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$, (d) stands for dummy variables. The social network variable is 'Relatives': it is the average number of relatives inside the village each farmer can ask for help or information.

Table 4.11: Social interactions involving less visibility, details of the regressions results

	Zero-Inflated Estimates: Number of people asked by each farmer:			*Logit Estimates*:
	for help in kind or in cash	About best farm practices	About market (output or land)	Asking for help on the experimental plot: planting or weeding
		Poisson model		*Logit model*
Improved seeds (d)	-0.05	0.25*	0.41*	0.31
	(0.14)	(0.13)	(0.22)	(0.21)
Relatives inside the village	0.18***	0.16***	0.16***	0.23***
	(0.06)	(0.04)	(0.03)	(0.07)
Improved seeds (d)*Relatives inside the village	-0.01	-0.06	-0.05	-0.06
	(0.07)	(0.06)	(0.06)	(0.08)
North	-0.11	0.11	-0.12	2.49***
	(0.18)	(0.21)	(0.39)	(0.39)
Been asked (d)	0.45***	0.49**	0.35	
	(0.17)	(0.20)	(0.37)	
Plot size (ha)	1.46	-2.65***	-2.29	2.17*
	(1.43)	(1.00)	(2.22)	(1.32)
Farm size (ha)	-0.04	0.03	0.15*	0.10
	(0.06)	(0.05)	(0.09)	(0.09)
Oxen (d)	0.24*	0.28***	0.04	-0.16
	(0.14)	(0.07)	(0.22)	(0.31)
Labour (man day)	0.01	0.00	0.01	
	(0.01)	(0.01)	(0.02)	
Pest damage (d)	-0.33***	-0.29*	-0.57***	0.99
	(0.12)	(0.15)	(0.15)	(0.66)
Standardized Precipitation Index	0.05	-0.26***	0.14	-0.51
	(0.12)	(0.07)	(0.17)	(0.37)
Female headed household (d)	0.35*	-0.02	-0.08	-0.69*
	(0.20)	(0.17)	(0.17)	(0.37)
Age	0.00	0.01	0.00	0.00
	(0.00)	(0.01)	(0.00)	(0.01)
Household size	0.04	-0.06**	-0.03	-0.12***
	(0.03)	(0.03)	(0.06)	(0.04)
Secondary education (d)	-0.08	0.12	-0.13	-0.08
	(0.13)	(0.16)	(0.22)	(0.35)
Risk averse (d)	0.16	0.09	-0.05	-0.05
	(0.20)	(0.15)	(0.16)	(0.37)
Leadership role in the community (d)	0.09	0.01	0.06	-0.01
	(0.10)	(0.10)	(0.12)	(0.32)
Member of a self-help group(d)	0.21*	0.32**	0.08	-0.01
	(0.13)	(0.14)	(0.27)	(0.33)
Member of a religious association(d)	0.08	-0.14	0.12	0.59**
	(0.11)	(0.10)	(0.18)	(0.25)
Walking distance to the plot (minutes)	0.01***	-0.00	0.00	0.02***
	(0.00)	(0.00)	(0.01)	(0.01)
Constant	0.34	0.61	0.78*	-3.36***
	(0.41)	(0.42)	(0.40)	(0.51)

Log odd of asking no one (logit model)

Improved seeds (d)	0.22	0.64	0.52	
	(0.38)	(0.94)	(0.47)	
Relatives inside the village	0.06	-0.99	-0.33**	
	(0.19)	(0.98)	(0.13)	
Improved seeds (d)*Relatives inside the village	-0.18	0.06	0.04	
	(0.25)	(1.31)	(0.16)	
North	0.08	0.63	0.81**	
	(0.39)	(1.10)	(0.36)	
Been asked (d)	-3.62***	-4.63***	-3.95***	
	(0.49)	(0.70)	(0.45)	
Farm size (ha)	0.07	-0.41	0.23	
	(0.18)	(0.52)	(0.19)	
Age	-0.01	-0.01	0.01	
	(0.01)	(0.03)	(0.02)	
Household size	0.19***	-0.15	-0.03	
	(0.06)	(0.21)	(0.11)	
Leadership role in the community (d)	0.47	0.27	-0.15	
	(0.35)	(0.69)	(0.30)	
Female headed household (d)	0.06	-0.91	0.23	
	(0.65)	(1.04)	(0.41)	
Secondary education (d)	-0.37	1.65***	0.37	
	(0.38)	(0.43)	(0.48)	
Risk averse (d)	0.87***	0.70	-0.01	
	(0.33)	(0.82)	(0.52)	
Constant	0.37	0.95	0.31	
	(0.95)	(1.07)	(0.75)	
Observations	313	313	313	313
McFadden's R^2	0.23	0.20	0.25	0.14
Wald-test	160.72***	101.13***	188.5***	47.47***
AIC	1543	1175	1154	374
BIC	1599	1231	1210	430

*Cluster robust standard errors at the village level in parentheses, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$, (d) stands for dummy variables. The social network variable is 'Relatives': it is the average number of relatives inside the village each farmer can ask for help or information.*

5 Agri-environmental Schemes and Risk exposure: a case study from Ireland

Xavier Vollenweider[1]

Abstract

We estimate the impact of an Irish agri-environmental scheme (AES) on farmer's risk exposure with the moment-based approach of Antle (1983) applied to a panel dataset covering the 2006-2009 period. The AES does not have a large impact on risk exposure; it even slightly decreases the variance of the net gross margin distribution. We then compute the risk premium across farm categories and find that the greatest benefit in terms of risk premium reduction goes to sheep farmers. The benefit of joining the AES is mostly driven by an increase in the expected gross margin. The dairy sector, the most intensive sector of Irish agriculture, is also the one which benefits the least from the scheme. This is consistent with the observation that dairy farmers are under-represented in the scheme.

[1] Department of Geography and Environment, London School of Economics and Political Science, United Kingdom. Contact author: Xavier Vollenweider: X.Y.Vollenweider@lse.ac.uk. I gratefully acknowledge the support from the Walsh fellowship and the Rural Economy and Development Programme of Teagasc, the Agriculture and Food Development Authority of the Republic of Ireland. I would like also to extend my thanks to Steve Gibbons, Cathal O'Donoghue and Salvatore Di Falco for comments on earlier drafts of this chapter.

5.1 Introduction

We present here a set of evidence of the impact of agri-environmental schemes (AES) on farmer risk exposure. The republic of Ireland provides an interesting case study as its AES, the Rural Environment Protection Scheme (REPS), is universal - i.e. all farmers can enter the scheme, not only the ones located in an environmentally sensitive area - and is voluntary (Emerson and Gillmor 1999). The goal of REPS is to incentivize farmers with financial rewards to adopt environmentally friendly practices over a five-year period. The aim of the present study is to estimate the impact of joining REPS on farmers' risk exposure. Several factors could explain the link between AES and risk exposure.

First, some production standards imposed by AES, such as a reduction in fertiliser and pesticide application rates, might have a direct impact on risk exposure. Organic and low input farmers have indeed been shown to be more exposed to production risk than conventional farmers (Berentsen et al. 2012; Finger 2014; Gardebroek 2006; Serra et al. 2008). We might therefore expected REPS farmers to be more exposed to production risk than those staying out of the scheme.

Second, Morris et al. (2000) report a concern among English farmers that the rigidity of the AES reduced their ability to take remedial action in case of pest infestation or severe weed events. AES contract length was found to negatively affect the decision to join because it tied farmers' hands over a long period of time (Peerlings and Polman 2009). According to these findings, joining an AES would increase risk exposure.

Third, the change in agricultural practices required to take part to an AES is generally perceived as a risk that younger farmers were more willing to take (Wynn et al. 2001). Corroborating the increase in risk linked to non-conventional farming, organic farmers have been shown to be less risk averse than non-organic farmers (Gardebroek 2006; Serra et al. 2008). It is however yet unclear if joining an AES is objectively riskier or if it is perceived as such by farmers because of the uncertainty linked to changes in long established farming practices. Our aim is to test the impact of REPS on farmers' risk exposure.

As the entire farming population of Ireland is eligible for REPS, the analysis is performed across all the major farm categories in Ireland. Our impact analysis is based on the moment-based approach of Antle (1983). REPS is introduced in section 5.2. We present the model and the estimation strategy in section 5.3 and 5.4. The data are presented in section 5.5 and the empirical implementation in section 5.6. Results follow in section 5.7 and we conclude in section 5.8.

5.2 The Rural Environment Protection Scheme

European Council Regulation 2078/92 required member states to implement policies fostering the adoption of environmentally sustainable agricultural production practices. The response of the Irish government was the design, in consultation with farming and environmental groups, of the Rural Environment Protection Scheme (REPS, Emerson and Gillmor 1999). REPS came into operation on the 1st of June 1994. Its three objectives are to incentivize farmers to: (1) produce food in an extensive and environmentally friendly manner; (2) to protect biodiversity, endangered species and wildlife habitat and (3) to preserve the landscape (DAFF 2005; Emerson and Gillmor 1999).

In order to benefit from the REPS subsidy, farmers have to draw up a five-year production plan so as to implement a comprehensive set of eleven mandatory measures, extending from waste management, fertilizer use and stocking rate, to the protection of wildlife habitats, historical remains and the improvement of the visual appearance of the farm (Emerson and Gillmor 1999). Over the period 1994 to 2009 - the year of its closure to new entrants -, there were four successive reforms of REPS. There has not been major change in the overall design of the policy, although additional payments for farms above 40 hectares (ha) and other supplementary measures were introduced in 2005. The goal of the scheme was thus to reward financially, over five years, the more environmentally virtuous farmer. Noncompliance with the agreed-upon farm management plan could lead to fines and ultimately to the exclusion from the scheme.

The first version of REPS, REPS I, reached its target of 45’000 farmers enrolled. It did not, however, attract the most intensive and polluting farmers (Murphy et al. 2014). Farmers were dissatisfied with REPS II as payments were considered too low and the administrative burden rate too high (Murphy et al. 2014). As a result, the target of 70’000 farmers was not attained, with just over 30’000 farmers participating. Nevertheless, the participation of dairy farmers (who tend to operate more intensive and polluting farms) increased. These farmers might have used REPS II as a risk management instrument following the loss of income caused by the foot-and-mouth disease epidemic in 2001 (Murphy et al. 2014). The uptake of REPS III, initiated in 2005, attracted many more farmers than earlier versions. Murphy et al. (2014) report that the Irish Farmers Association endorsed the new REPS even before it was introduced while contemporary newspapers relayed a sense of urgency to enrol before the program closed to new entrants.

In REPS III, two additional biodiversity measures had to be adopted out of a choice of 16 possible measures. Subject to the fulfilment of their action plan, farmers received: 200 euros per hectare for the first 20 ha;

175 euros per ha for the next 20 ha up to 40 ha; 70 euros/ha for the next 15 ha up to 55 ha and 10 euros per ha for areas over 55 ha. The last two payment tiers were a novelty in REPS III and contributed to its appeal. Higher payments were given to environmentally sensitive areas and supplementary measures such as organic farming practices lead to additional payments.

The introduction of the Single farm payment in 2005 and the ratification of the Nitrate directive in 2006 may have reduced the opportunity cost of joining REPS. Indeed, some of the accompanying measures, mandatory for receiving the Single Farm Payment, were already present in REPS (Murphy et al. 2014). Furthermore, the decoupling of subsidy from output level might have reduced the incentive to overproduce, favouring production plans more compatible with REPS (O'donoghue and Howley 2012). In REPS IV, applied from 2008 until 2009, farmers using more than 170kg of nitrogen per ha could apply for a derogation under the Nitrate directive. The goal was to further increase the participation of intensive farmers. Furthermore, farmers could choose among a larger option of biodiversity conservation measures.

In the present study we focus on the period 2006 to 2009. Most farmers were under the REPS III contract over this period. REPS payments were relatively stable over the years covered in our study, with an average of 6500 euros per participating farmer. Almost half of the farms in the sample were taking part in the scheme.

5.3 Theoretical background

Our goal is to estimate the impact of REPS on farmers’ risk exposure by investigating the conditional distribution of their net gross margin. Net gross margin is defined as the difference between gross output and direct costs net of REPS subsidy. Gross output might be affected by weather, pests or diseases and other random factors, while prices swings might affect input costs and output value. The net gross margin can hence be expressed for each farm by a conditional distribution function:

$$F(\pi|\boldsymbol{\mu_{it}}) \quad i = 1, \dots, N \qquad t = 1, \dots, T \tag{1}$$

where π is the net gross margin defined as a random variable and $\boldsymbol{\mu_{it}} = (\mu_{1it}, \dots, \mu_{mit})$ is a vector of the m central moments characterizing the net gross-margin distribution for farmer i at time t (Gardebroek 2006). The conditional distribution function is thus assumed to be the same for all farmers, but the moments of each farmer’s conditional distribution are allowed to differ between farmers, and are related to farm characteristics and input choices. The first moment is the expected net gross margin. The second moment is the variance of net gross margin, and gives a sense of the spread of the net gross margin a farmer can expect. The third moment captures the asymmetry of the distribution of the possible gross-margins, negative values implying the presence of downside risk. The econometric strategy to estimate these central moments is presented in the next section.

In order to assess the riskiness of each farmer's conditional distribution, we rely on two indicators summarizing these moments. The first one is the coefficient of variation, defined as the standard deviation of the conditional distribution divided by its mean, $\sqrt{\mu_{2it}}/\mu_{1it}$.

The second indicator is the risk premium, R_{it}. It can be interpreted as the implicit cost of risk bearing, i.e. the maximum price one is ready to pay to get rid of all risk. Following Pratt (1964), the risk premium satisfies:

$$EU(\pi_{it}) = U[E(\pi_{it}) - R_{it}] \tag{2}$$

where *E* stands for the expectation operator, and U is a utility function we will define below.

Rearranging and approximating the last expression by its Taylor expansion of degree three (e.g. John M Antle 1987; Chavas and Holt 1996b), we can express the risk premium as:

$$R_{it} \approx \frac{1}{2} AP\mu_{2it} \ - \frac{1}{6} DS\mu_{3it} \tag{3}$$

where μ_{mit} is the m central moment for $m = 2{,}3$ of farmer i at time t, AP is the coefficient of absolute risk aversion (Pratt 1964) for mean-preserving spread aversion, and DS is the coefficient of downside risk aversion (Menezes et al. 1980) for mean-spread-preserving skewness preferences. The risk premium depends on:

- the set of risk preference parameters, $AP$and DS;
- the variance of the conditional net gross margin distribution of each farmer, μ_{2it};
- the third central moment of the conditional net gross margin distribution of each farmer, μ_{3it}.

Several models have been applied for the estimation of risk preferences, either based on recursive estimation (John M Antle 1987; 2010; Foudi and Erdlenbruch 2011; Groom et al. 2008; Simtowe et al. 2006), or on joint estimation of preferences and technology parameters (Chavas and Holt 1996b; Koundouri et al. 2009; S. C. Kumbhakar 2002; S. C. Kumbhakar and Tveterås 2003; S. Kumbhakar and Tsionas 2010; Love and Buccola 1991; Pope and Just 1991; Saha et al. 1994; Saha 1997; Vollenweider et al. 2011). As regards the structure of risk aversion, the results point toward declining absolute risk aversion (Bar-Shira et al. 1997; Chavas and Holt 1990, 1996b; Saha et al. 1994), and increasing relative risk aversion (Bar-Shira et al. 1997; Saha et al. 1994).

Lence (2009) calls into question the validity of estimating risk aversion with agricultural production data, particularly when shocks are not large or the sample is small. Just and Just (2010) show that although the models can be locally identified, they are not globally identified: an infinite set of pairs of technology and utility functions can equally well fit the data.

These last two papers are a serious blow to the field of risk preference estimation with production data. Hence, we use a more straightforward approach whereby we assume a given parametric form for the utility function and present the risk premium according to the various degrees of risk aversion commonly found in the literature.

Following the literature on risk preference estimation, farmers are assumed to exhibit Declining Absolute Risk Aversion (DARA). Thus, we model the utility function with a power utility function:

$$U(x) = \frac{x^{1-\gamma}}{(1-\gamma)} \tag{4}$$

where γ is the coefficient of relative risk aversion. Farmers hence exhibit DARA and constant relative risk aversion. Although it has been found that the degree of risk aversion differs between farmers, notably in the case of organic vs. conventional farmers (Gardebroek 2006; Serra et al. 2008), we will assume that the overall structure of risk preferences and the parameter of relative risk aversion remain unchanged in and out of the scheme.[2] Nevertheless, as farmers exhibit DARA, an increase in expected net gross margin decreases their level of absolute risk aversion. We present below the estimation strategy for recovering the central moment of each farmer's conditional net gross margin distribution.

[2] Nauges et al. show that changes in subsidy policies might have an impact on producers risk preferences (Kounduri et al., 2009). In a different context, an increase in violence linked to civil conflict in Burundi have been shown to decrease risk aversion (Voors et al., 2012).

5.4 Estimation Strategy

We rely on the 'moment-based' approach of Antle (1987) to estimate the expected net gross margin, its variance and skewness as defined in equation (1) and the marginal effect of REPS on them. We follow below closely his derivation (see also Gardebroek 2006).

Let us define the probability density of net gross margin as $f(\pi|\boldsymbol{x_{it}})$ where π is the stochastic net gross margin and $\boldsymbol{x_{it}}$ is a vector of farmer's i input and characteristics at time t. The moments of the net gross margin can be written in general form as (John M Antle 1987):

$$\mu_{it1}(\boldsymbol{x_{it}}) = \int \pi\, f(\pi|\boldsymbol{x_{it}})\mathrm{d}\pi \tag{5}$$

$$\mu_{itm} = \int \left(\pi - \mu_{it1}(\boldsymbol{x_{it}})\right)^m f(\pi|\boldsymbol{x})d\pi \tag{6}$$

where $\mu_{it1}(\boldsymbol{x_{it}})$ is the first moment, and $\mu_{itm}(\boldsymbol{x_{it}})$ are the m moments for $m \geq 2$. The moments are hence a function of a vector of inputs and farmer characteristics, $\boldsymbol{x_{it}}$. Assuming a linear relationship between the moment and the variables, both equations in (5) and (6) can be expressed as:

$$\mu_{it1}(\boldsymbol{x_{it}}) = \boldsymbol{x_{it}\beta_1} \tag{7}$$

$$\mu_{itm}(\boldsymbol{x_{it}}) = \boldsymbol{x_{it}\beta_m} \tag{8}$$

We will test in the empirical section several specifications for $\boldsymbol{x_{it}\beta_1}$ and $\boldsymbol{x_{it}\beta_m}$. Antle (1983) and Gardebroek (2006) use a quadratic equation. As the net gross margin is random, we can write the first moment as the following regression equation:

$$\pi_{it} = \boldsymbol{x_{it}\beta_1} + v_{it} \tag{9}$$

where π_{it} is the net gross margin of farmer i at time t, $\boldsymbol{x_{it}\beta_1}$ is the expected net gross margin at time t and v_{it} an error term with expectation equal to zero. Based on equation (9), we have $E[(\pi_{it} - \mu_{it1})^m] = E({v_i}^m) \equiv \mu_{itm}$, for $m \geq 2$.

The regression equation for the m central moment is given by:

$$(v_{it})^m = \boldsymbol{x_{it}\beta_m} + v_i \tag{10}$$

Letting $\boldsymbol{x_{it}}$ contain a binary variable equal to one for farmers having joined REPS and zero otherwise, we can estimate the impact of REPS on expected net gross margin, variance and skewness. Furthermore, we can use the estimated moment function in equation (9) and (10) as a building block for the risk premium given in equation (3). Note that equation (10) corresponds to an assumption that there is heteroskedasticity in the moment equations. The estimates of $\boldsymbol{\beta_1}$ and $\boldsymbol{\beta_m}$ are hence not efficient, but they are still consistent. Therefore the difference in variances and skewness of net gross margin between REPS and non-REPS farmers are estimated consistently (Gardebroek 2006).

5.5 Data

We rely on data from the Irish National Farm Survey to conduct the analysis. This is an annual survey collected by Teagasc, a semi-state research body of the Republic of Ireland and feeding into the European Farm Accountancy Data Network (FADN). We selected the 2006 to 2009 years because the system of subsidies is comparable over this period. Indeed, the 2005 Common Agricultural Policy reform unified a large part of Pillar I farm support under the single farm payment. Furthermore, most farmers over the period were part of the third version of REPS III (Murphy et al. 2014). The number of farms in the dataset oscillates between 860 and 1052 per annum, and farms took part to the survey 3 years in a row on average. Farming activities are divided into 6 categories according to the FADN convention. The sample is composed of specialist dairy (17%), dairy and other (8%), cattle (21%), cattle and other (23 %), sheep (11%), tillage (9%).

We use as dependent variable the farm net gross margin, i.e. the difference between gross output and direct costs minus the REPS subsidy. Gross output is the sum of the livestock gross output, the crop gross output, farm subsidies, other farm income as well as inter enterprise transfers[3]. Direct costs include the purchased feed concentrate, bulky feed, hired casual labour, the value of the fertilisers and pesticide (or other crop protection) used, machinery hire cost, veterinary costs and other miscellaneous expenses.

The explanatory variables are the capital expenditure (capital expenditure during the year less capital sales and capital grants[4]), land (the utilised agricultural area of the farm in hectares), labour (total number of

[3] The gross output is 'total sales less purchase of livestock, plus value of farm produce used in the house plus receipts for hire work, service fees etc. It also includes net change in inventory which for cows, cattle and sheep is calculated as the change in numbers valued at closing inventory prices. All non-capital grants, subsidies, premia, headage payments etc are also included, as are income from land and quota let' (RERC, 2009).

[4] Major repairs to farm buildings, plant and machinery and land improvements are included.

labour units working on the farm[5]), a diversification index (Berry 1971) based on the share of each source of gross margin within total gross margin[6], a series of dummy variables on main farming categories (Dairy, Cattle, Sheep, Tillage). Farms are classified according to dominant enterprise. As farms often have multiple enterprises, for instance cattle and dairy, individual farmers might change categories over the years.

Table 5.1: Variables used in the analysis

	REPS farmers	Non-REPS farmers	Difference
Gross margin	43,685.35	40,640.30	3,045.04***
	(26,489.07)	(31,662.43)	(977.91)
Gross margin net of REPS subsidies	37,549.90	39,828.20	-2,278.29**
	(26,229.03)	(30,861.61)	(958.87)
REPS subsidy	6,452.45	0.00	6,452.45***
	(2,066.36)	(0.00)	(47.78)
Capital	8,280.43	8,321.75	-41.32
	(14,396.34)	(15,344.23)	(498.75)
Land	45.04	43.28	1.76**
	(20.92)	(23.58)	(0.75)
Labour	1.13	1.17	-0.04***
	(0.40)	(0.44)	(0.01)
Berry diversification index	0.43	0.46	-0.03***
	(15.63)	(14.11)	(0.50)
Cattle (d)	0.50	0.44	0.06***
	(0.50)	(0.50)	(0.02)
Tillage (d)	0.09	0.08	0.01
	(0.29)	(0.27)	(0.01)
Sheep (d)	0.15	0.08	0.07***
	(0.36)	(0.27)	(0.01)
Dairy(d)	0.26	0.41	-0.14***
	(0.44)	(0.49)	(0.02)
Observations	1,907	2,066	3,913

Standard deviation in parentheses in column 1 and 2, standard error of the t-test of comparison of means in parentheses in column 3, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

[5] 1'800 hours per year is worth one labour unit, but one person cannot work more than one labour unit even if he/she works more than 1'800 hours per year. People under 18 years of age are given the following labour-unit equivalent: 16-18 years = 0.75; 14-16 years = 0.50.

[6] Gross margin sources are classified in 9 categories: Dairying, cattle, sheep, pigs, poultry, horses, crops, hire of machinery revenue, other current receipts. The Berry diversification index is expressed as:

$$B = 1 - \sum_{k=1}^{9} s_k^2$$

where s_k^2 is the share of each source k in gross margin.

Table 5.1 presents summary statistics over the whole period (2006-2009). The third column presents the average difference between REPS and non REPS participants. REPS farmers tend to have a higher gross margin than non-REPS farmers, however the difference is mainly driven by REPS subsidy as REPS farmers have a lower *net* gross margin. There are no differences in labour or capital between both categories.

The literature show contrasting evidence on the role of farm size on the decision to join an AES. Some studies finds a strong positive role (e.g. Damianos and Giannakopoulos 2002; Mazorra 2001), others a negative one (e.g. Vanslembrouck et al. 2002), while still other find no impact (Wynn et al. 2001). This heterogeneity in findings might be due to the coexistence of large extensive farms and intensive farms, the former having a lower opportunity cost to join the AES as suggested by Murphy et al. (2014) because their default farm management practices are more in accordance with REPS requirements. In our sample, landholding is slightly larger on REPS farms.

The largest differences between REPS and non REPS farmers are in terms of shares of sheep and dairy farms. The share of sheep farms is higher in the REPS category while the share of dairy farms is higher in the non-REPS category. This likely reflects the cost of joining REPS which tends to be lower for sheep farms and higher for dairy farms as the former tend to have extensive farming practices (<170 kg of nitrogen per ha) and the latter intensive farming practices (>170 kg of nitrogen per ha). In other terms, the default farm management of sheep farms are in accordance with REPS.

Table 5.2: Socio-economic variables and farm characteristics

Variable	REPS farmers	Non-REPS farmers	Difference
Farm family income	24,701.79	22,138.24	2,563.55***
	(20,351.70)	(23,572.17)	(719.72)
Farm family income per capita	13,889.65	12,709.55	1,180.11***
	(11,386.32)	(13,589.90)	(409.98)
Age between 25 and 44	0.63	0.55	0.08***
	(0.83)	(0.80)	(0.03)
Age above 65	0.40	0.54	-0.14***
	(0.66)	(0.75)	(0.02)
Number of household members with third	0.20	0.17	0.03*
level education	(0.53)	(0.48)	(0.02)
Good soil	0.49	0.53	-0.04**
	(0.50)	(0.50)	(0.02)
Bad soil	0.11	0.09	0.02*
	(0.32)	(0.29)	(0.01)
Herd size on dairy farms	39.26	47.76	-8.51***

	(22.20)	(25.16)	(1.39)
Observations	1,907	2,006	3913

Standard deviation in parentheses in column 1 and 2, standard error of the t-test of comparison of means in parentheses in column 3, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

We present in Table 5.2 a series of socio-demographic and farm characteristics variables. The farm family income, i.e. gross margin minus over-head cost[7], is higher among REPS participants. They also tend to be younger. This is coherent with the fact that younger farmers tend to have a higher likelihood of joining AES because they are ready to take the risk of adopting the new farming practices required in the AES (Wynn et al. 2001). The average number of household member with a third level education is also higher, in line with Dupraz et al. (2003) who showed that education generally encourages participation in AES. REPS participants tend to have lower quality soils, which might reflect their lower opportunity cost in joining REPS as intensive farming is less profitable on poor quality soils (Hynes and Garvey 2009). Similarly, herd size on REPS dairy farms are smaller, suggesting that smaller scale exploitations joins REPS while larger one stay out, pointing toward an adverse selection process in REPS as suggested by Hynes and Garvey (2009).

5.6 Empirical implementation

The moments of net gross margin presented in equations (9) and (10) are estimated sequentially. Although sequential estimation is not the most efficient approach, it is consistent. It was originally applied by Antle (1983) and more recently by Kim and Chavas (2003), Gardebroek (2006), Di Falco and Chavas (2006, 2009), Bangwayo-Skeete et al. (2012) and by Gardebroek et al. (2010) to the related Just and Pope framework for production risk estimation.

Outliers are identified as observations with a net gross margin greater than the yearly farm category average by 3 standard deviations (1% of the sample). The explicative variables are labour, capital, land, a diversity index, a series of dummy variables for the three farm categories (Cattle, Sheep, Tillage; Dairy being the base category), a dummy for REPS participation as well as a series of time fixed effects. We also added a series of interaction terms between the REPS dummy and the farm category dummies in order to let the impact of REPS on expected net gross margin and higher moments vary across farm categories.

[7] Over-head cost include cost of land rental, car, electricity, telephone, interest payment, depreciation of the machinery, machinery operating expenses, depreciation of the building, building repairs and upkeep, depreciation of land improvement work, land general upkeep, accountants and other consultant's fees and other miscellaneous expenses.

Table 5.3: Specification tests

		AIC	BIC
Expected net gross margin, μ_{it1}	Quadratic	83996.78	84128.81
	Quadratic without interaction	84028.64	84122.95
	Linear	84028.64	84122.95
Variance, μ_{it2}	Quadratic	161244.7	161370.4
	Quadratic without interaction	161293.8	161407
	Linear	161290.9	161385.2
Skewness, μ_{it3}	Quadratic	104668.6	104792.3
	Quadratic without interaction	104662.8	104767.9
	Linear	104658.4	104744.9

For the labour, capital and land variables, we tested three functional forms: a quadratic, a quadratic without interaction, and a specification where the variables enter only in level. Table 5.3 shows the results of the Bayesian and Akaike information criteria tests (respectively BIC and AIC) for the three first moments. For the first central moment, the quadratic is preferred according to the AIC while, according to the BIC, the quadratic functional form without interaction is better. In order to minimize mis-specification errors, which could cascade across the whole model via the use of the residual of the first moment equation for the estimation of the second and third moments equations (see section 5.4), we privileged results given by the AIC as it penalizes model complexity less compared to model fit. Furthermore, the quadratic form provides a Taylor second order approximation of any unknown expected net gross margin function (e.g. S. C. Kumbhakar and Tveterås 2003) and should thus, on average, produce a better fit to the data than the quadratic without interaction. AIC and BIC results agree for the second and third moments: the quadratic is preferred for the second moment and the linear specification for the third. The specifications of the first three moments are given in equations (11) to (13):

$$\begin{aligned}\pi_{it} = \beta_{01} + \sum_{j=1}^{3}\beta_{j1}x_{jit} + \frac{1}{2}\sum_{j=1}^{3}\sum_{k=1}^{3}\beta_{jk1}x_{jit}x_{kit} + \beta_{R1}REPS_{it} + \sum_{c=1}^{3}\beta_{c1}c_{cit} \\ + \sum_{c=1}^{3}\beta_{Rc1}c_{cit}REPS_{it} + \beta_{d1}Div_{it} + \sum_{h=1}^{3}\beta_{h1}T_{hit} + \alpha_{1i} + v_{1it}\end{aligned} \tag{11}$$

$$\begin{aligned}(\hat{v}_{1it})^2 = \beta_{02} + \sum_{j=1}^{3}\beta_{j2}x_{jit} + \frac{1}{2}\sum_{j=1}^{3}\sum_{k=1}^{3}\beta_{jk2}x_{jit}x_{kit} + \beta_{R2}REPS_{it} + \sum_{c=1}^{3}\beta_{c2}c_{cit} \\ + \sum_{c=1}^{3}\beta_{Rc2}c_{cit}REPS_{it} + \beta_{d2}Div_{it} + \sum_{h=1}^{3}\beta_{h2}T_{hit} + \alpha_{2i} + v_{2it}\end{aligned} \tag{12}$$

$$(\hat{v}_{1it})^3 = \beta_{03} + \sum_{j=1}^{3} \beta_{j3} x_{jit} + \beta_{R3} REPS_{it} + \sum_{c=1}^{3} \beta_{c3} c_{cit} + \sum_{c=1}^{3} \beta_{Rc3} c_{cit} REPS_{it} + \beta_{d3} Div_{it} + \sum_{h=1}^{3} \beta_{h3} T_{hit} + \alpha_{3i} + v_{3it} \tag{13}$$

where x_{jit} is land, capital and labour for $j = 1,2,3$ respectively, $REPS_{it}$ is a dummy equal to one if the farmer is a REPS participant, c_{cit} is a binary for the sheep, cattle and tillage categories for the subscript $c = 1,2,3$ respectively (the base category is dairy), Div_{it} is the Berry diversification index, T_{hit} are time-fixed effects for $h = 2007,2008,2009$, α_{mi} are the farm fixed effects and v_{mit} an error term for moments $m = 1,2,3$.

Equations (10) to (12) are estimated with a fixed effect estimator (within effect). We tested with a Durbin-Wu-Watson if REPS was endogenous and failed to reject the null of no endogeneity (p-value= 0.325). We did find some correlation in the error terms, but as the panel is short, this should not affect too largely the results. We use robust standard error to clustering at the farm level.

5.7 Results

Table 5.8 in the appendix shows the results of the estimation of the first moment of net gross margin and Table 5.4 shows the marginal effects. Out of the 9 variables of the quadratic function of labour, land and capital, three are found to be significant. We nevertheless chose to keep all the variables of the quadratic function (i.e. land, labour and capital in level, square and their interactions) in order to conserve flexibility in the expected net gross margin function.[8] The other estimated parameters are very significant and the R-squared is reasonably high: 30% of the within variation is explained by the model.

Table 5.4 shows the marginal effects on the moments of farmers' net gross margin distributions. The effect of capital investment in machinery is minor as 1 euro invested increases net gross margin by only 2.5 cents. This low return on capital might reflect over-capitalisation of the farms. Land, by contrast, has a positive

[8] Kumbhakar and Tveterås (2003) estimated a related model of production risk with a quadratic function. They notes: '[t]he idea of dropping insignificant variables is not pursued […] due to several problems. First, it destroys the flexibility of the mean output function. Second, dropping one insignificant variable caused other insignificant (significant) variables to be significant (insignificant) due to high multicollinearity (which is always present in flexible functions) and the use of a system approach. Furthermore, we found no natural order to select variables for exclusion in the present model'.

effect on net gross margin; an additional hectare increase net gross margin by 160 euros. Diversification also increases net gross margin. All the farm category dummies are negative as expected because the base category, the dairy sector, is the most profitable sector of the Irish agriculture. Switching from dairy production to another system is therefore on average not profitable.

The average effect of REPS on expected net gross margin is negative: farmers joining REPS lose on average 3'000 euros. This might be due to an increase in the direct costs incurred to comply with the REPS plan as well lower output caused by the reduction in fertiliser use and stocking rate. Nevertheless, the cost of joining REPS appears to be more than compensated by the REPS subsidy (7'060 euros on average). Hence REPS III provided more than adequate compensation for possible revenue loss and the increase in cost due to compliance with the REPS production plan. This likely explains the large success in the uptake rate of REPS III.

The impact of REPS differs however across categories. Dairy farms are those who benefit the least from REPS; it decreases their expected net gross margin by 7'000 euros while the average subsidy in the dairy category is 7'500 euros. As dairy farms are the most productive and intensive farm category, production constraints (e.g. reduction in stocking rate and fertiliser use) might cause higher compliance costs and revenue losses from REPS participation. This low average net benefit of joining REPS (circa 500 euros) is consistent with the observed low participation rate of dairy farmers. Being tied up over 5 years in farm management plan as well as the administrative burden of joining REPS might not have been adequately compensated by the REPS subsidy.

By contrast, sheep farms, the most extensive production system, is the one benefits greatly from joining the scheme. Joining REPS causes a drop of 4’000 euros in expected net gross-margin of sheep farmers, but it is largely compensated by the REPS subsidy equal, on average, to 7’850 euros. As they are more likely to meet limitations in nitrogen usage by default, joining REPS implies a lower cost than for other farm categories. As noted by Hynes and Garvey (2009), the design of REPS might have created an adverse selection whereby farms for which REPS measures do not constitute a large change in their farm management join, while the more intensive and polluting ones, i..e the dairy farms, do not. Murphy et al. (2014) report that sheep farms are over-represented in the REPS compared to the share of sheep farms in Irish agriculture. The effect of joining REPS on cattle and tillage farms does not differ significantly from the sample average, i.e. it decreases gross margin by 3’000 euros while the average REPS subsidy in both sector is, respectively, 6’400 and 7’850 euros.

Table 5.4: Marginal effects on the moment of net gross margin

Marginal effect	Mean	Variance
capital	0.025*	-142.829
	(0.014)	(141.421)
labour	-1,029.500	755,921.037
	(1,823.398)	(18,822,286.615)
land	158.226**	523,155.405
	(68.629)	(758,625.301)
diversification	59.068**	19,883.734
	(26.737)	(178,561.798)
cattle farms	-5,038.581***	-57,284,248.785***
	(1,873.306)	(22,280,288.567)
tillage farms	-6,294.716*	-87,557,742.616***
	(3,383.684)	(31,226,220.979)
sheep farms	-4,063.329**	-52,363,719.018**
	(1,730.428)	(20,891,646.485)
REPS average	-2997.076***	-14782600*
	(922)	(8400618)
REPS for dairy farms	-6,789.095***	-41,505,350.737**
	(1,760.024)	(18,093,153.704)
REPS for cattle farms	-544.90	9689194.1
	(1165.9)	(7459878)
REPS for sheep farms	4110.354**	43167074***
	(1935.625)	(14308620)
REPS for tillage farms	-1811.099	-10687388
	(1935.625)	(14308620)

* *p < 0.10,* ** *p < 0.05,* *** *p < 0.01, standard errors in parentheses*

Table 5.8 in the appendix shows the regression results for the second central moment defined in equation (12). Three of the nine variables in the quadratic function of labour, land and capital are significant while all the other variables are significant. The r-squared is 7%, a figure in line with other studies (e.g. Bangwayo-Skeete et al. 2012; Gardebroek et al. 2010).

Table 5.4 presents the marginal effect for the second moment equation. Capital, labour and land have no effect on variance while shifting from the dairy sector to another sector decreases net gross margin variance. REPS is found to be risk-diminishing on average. At the farm sector level, REPS decreases variance for the dairy farms. The smaller variance for REPS farmers, although statistically significant, might however be driven by the smaller expected gross margin. We will investigate this further by computing the implicit cost of risk bearing for various degree of risk aversion.

Results for the third moment were not conclusive: the only parameters found to be significant are the year dummies and the diversification index. The latter increases skewness, i.e. it decreases the amount of downside risk as expected. Diversifying farm income sources across multiple enterprises might indeed help reduce the risk of a big loss caused by an adverse price swing of a single commodity, for instance. However, results of the third moment must be considered with care. Indeed, they rely heavily on the parametric assumption that the moments of the net gross margin distribution function can be modelled as a linear combination of variables. Furthermore, specification errors in the first moment equation are raised to the power three in the third moment equation, making the results very sensitive to any mis-specification in the first moment. Even though results are expected to be consistent, the number of observations required to achieve consistency may be very large. Along with Gardebroek (2006), we decided not to include the third moment in the computation of the risk premium presented below. Results for the third moment equation are nevertheless available in Table 5.8 in the appendix.

Table 5.5 displays the sample average central moment of farmers' net gross margin distributions as well as the results of t-test of comparison in means between REPS and non-REPS farmers. REPS farms have an average expected net gross margin 7'000 euros smaller than non-REPS farms (Table 5.1). Again, it suggests that less profitable farms tend to have a higher participation in REPS, underlying the role of REPS as an income support mechanism. REPS farmers also have a lower variance than non-REPS farmers on average, confirming the results found above. However, the coefficient of variation is only marginally smaller for REPS farmers (21% against 23%) and the difference is not statistically significant.

Table 5.5: Central moments of the net gross margin distribution

	REPS	Non-REPS	REPS-NON REPS
Expected net gross margin	40,748	47,444	-6,696***
	(30,360)	(38,325)	(1,126)
Variance	52,288,395	63,812,123	-11,523,728***
	(59,754,849)	(67,384,120)	(2,148,041)
Skewness	3,896,695,231	12,421,218,599	-8,524,523,368
	(202,870,556,290)	(201,048,547,121)	(6,576,589,135)
Coefficient of variation	0.2072	0.2283	-0.07
	(1.8623)	(0.1671)	(0.0064)

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$, *standard errors in parentheses*

Table 5.6 shows the sample average risk premium for REPS participants and non-participants according to various degree of risk aversion as well as the results of t-tests of means comparison. We do not consider skewness in the computation of the risk premium because of the lack of significant estimates in the third moment equation. Furthermore, the sample average skewness does not differ statistically between REPS and non-REPS farmers (see Table 5.5). The risk premium is therefore defined as:

$$R_{it} \equiv \frac{1}{2} AP\hat{\mu}_{2it} \qquad (14)$$

where $\hat{\mu}_{2it}$ is the variance of net gross margin at time t. The Arrow-Pratt coefficient of absolute risk aversion, AP, depends on the coefficient of relative risk aversion, γ. The literature on risk aversion estimation presents a wide range of estimates, from -0.10 in India (John M Antle 1989) to 7.62 in the US (Chavas and Holt 1996a). The most common values chosen in studies comparable to the present one oscillate between 2 and 3 (e.g. Di Falco and Chavas 2006; Finger 2014; Ligon and Schechter 2003), we present the results with γ ranging from 1 to 5 (risk premium is zero for $\gamma = 0$).

The risk premium is presented both in terms of *net* gross margin and gross margin with the REPS subsidy. In the latter case, we first add the value of the subsidy from REPS to the estimated expected net gross margin. We then compute the risk premium.

Table 5.6: Risk premium

	Sample average risk premium net of REPS subsidy			Sample average risk premium with REPS subsidy		
Risk aversion γ	REPS farmers	Non-REPS farmers	Difference	REPS farmers	Non-REPS farmers	Difference
1	717	757	-40^{*}	583	757	-174^{***}
	(660)	(647)	(23)	(518)	(647)	(20)
2	1,434	1,514	-80^{*}	1,166	1,514	-348^{***}
	(1,320)	(1,294)	(45)	(1,035)	(1,294)	(41)
3	2,151	2,271	-120^{*}	1,749	2,271	-521^{***}
	(1,980)	(1,940)	(68)	(1,553)	(1,940)	(61)
4	2,868	3,028	-160^{*}	2,333	3,028	-695^{***}
	(2,640)	(2,587)	(91)	(2,071)	(2,587)	(82)
5	3,585	3,784	-200^{*}	2,916	3,784	-869^{***}
	(3,300)	(3,234)	(113)	(2,589)	(3,234)	(102)

* *$p < 0.10$, ** $p < 0.05$, *** $p < 0.01$, standard errors in parentheses*

The differences reported in Table 5.6 show that the implicit cost of risk is lower for REPS farmers, although the difference is small. Once the REPS subsidy has been added to the expected net gross margin of REPS, the difference in risk premium is increases slightly, it is between 348 euros and 521 euro for moderate levels of risk aversion ($\gamma = 2{,}3$). The additional benefit of REPS in terms of risk reduction is hence quite low on average (5% to 7% of the average value of the REPS subsidy).

Table 5.7: Differences in risk premium between REPS and non REPS farmers in the different farming categories

	Sample average difference in risk premium net of REPS subsidy				Sample average difference in risk premium with REPS subsidy			
Risk aversion γ	Dairy	Cattle	Sheep	Tillage	Dairy	Cattle	Sheep	Tillage
1	-59^{**}	-56	-199^{*}	-86	-132^{***}	-237^{***}	-436^{***}	-163^{***}
	(30)	(47)	(120)	(64)	(29)	(42)	(101)	(58)
2	-118^{**}	-112	-398^{*}	-172	-264^{***}	-473^{***}	-872^{***}	-327^{***}
	(60)	(94)	(240)	(128)	(57)	(83)	(202)	(116)
3	-176^{**}	-168	-598^{*}	-258	-396^{***}	-710^{***}	$-1,308^{***}$	-490^{***}
	(89)	(141)	(360)	(191)	(86)	(125)	(304)	(174)
4	-235^{**}	-223	-797^{*}	-343	-528^{***}	-946^{***}	$-1,744^{***}$	-654^{***}
	(119)	(188)	(480)	(255)	(115)	(166)	(405)	(231)
5	-294^{**}	-279	-996^{*}	-429	-661^{***}	$-1,183^{***}$	$-2,181^{***}$	-817^{***}
	(149)	(235)	(601)	(319)	(144)	(208)	(506)	(289)

* *$p < 0.10$, ** $p < 0.05$, *** $p < 0.01$, standard errors in parentheses*

Table 5.7 shows the difference in risk premium between REPS participants and non-participants across farm categories. The greatest benefit in terms of risk reduction is in the sheep sector. For a degree of relative risk aversion equal to 3, the benefit amounts to 1'300 euros in the reduction of the implicit cost of risk bearing, i.e. an increase in the benefit from REPS of 16% compared to the sole subsidy (7'865 euros on average). By contrast, the additional benefit brought by REPS for dairy farmers represent only 5% of the amount of the REPS subsidy.

The low effect of REPS on risk exposure might be the result of the flexibility in the design of the scheme. Each farmer designed their own production plan in accordance with a farm adviser accredited by the government. This flexibility in the design of the policy rather than a one size-fits-all approach might have allowed farmers to design plans minimizing the impact of REPS on their risk exposure. Furthermore, most REPS participants are extensive farms on which no major changes are required in terms of fertiliser or pesticide use. Although we did find that REPS decreases the variance of the net gross margin distribution, there were no statistical differences in the coefficient of variation between participants and non-participants while differences in the risk premium computed with the net gross margin were economically very small. REPS had, therefore, mostly an impact on gross margin distribution via the increase in expected gross margin caused by the subsidy.

5.8 Conclusion

Studies on the impact of organic farming and low-input agriculture have shown that these production systems were riskier than conventional ones. Furthermore, the literature on adoption of agricultural environmental scheme (AES) has found that many farmers perceived agri-environmental schemes as risk-increasing. This does not appear to be the case for REPS as it slightly decreases, on average, both the variance of net gross margin and the implicit cost of risk bearing. This is true for all farm categories.

Most of the benefit in joining REPS is driven by an increase in expected gross margin. We found that joining REPS implied an average decrease in net gross margin of 3'000 euros. As the subsidy is on average 7'000 euros, the participants in REPS benefited from the scheme. REPS acted hence as a gross margin support subsidy.

However, the benefit of joining REPS varies between farm categories. Sheep farmers tend to benefit largely from the scheme while the gain is minimal for dairy farmers. The former ones benefit not only in terms of increased expected gross margin, they also benefit from a significant decrease in risk exposure. This might explain their over-representation in REPS compared to the other farm types. Hynes and Garvey (2009) showed that REPS suffered from adverse selection: the less polluting farmers participate (e.g. sheep farmers), while the most polluting farmers do not (e.g. the dairy farmers). As intensive farmers are generally the most profitable farms and face high compliance cost with AES measures, increasing their participation only via an increase in the subsidy might be hard to achieve and very expensive. An option would be to link AES with risk management schemes as proposed in the case of organic farmers by Serra et al. (2008). This could decrease the perceived riskiness of joining AES and increase the net benefit.

The moment-based approach adopted in the current paper provides a direct route for the estimation of all central moments of the distributions of gross margin at the farm level. We did not however find significant results in terms of the skewness of net gross margin. The lack of significant results for skewness might be caused by two factors. First, the sequential estimation strategy is sub-optimal in terms of efficiency. A maximum likelihood approach could increase the efficiency of the results. Furthermore, a dynamic approach could produce further insights on the impact of REPS on risk exposure. Indeed, the cost of joining an AES might vary over the duration of the AES contract while the implementation of AES measures is generally progressive over the overall length of the contract. It remains however to be shown that dynamic panel data methods are applicable to the methods of moments. This could constitute an interesting follow-up of this paper.

The second reason for a lack of significant results for the third moment is likely inherent to the use of the residuals of the first moment regression as dependent variables in the higher moments equations. Any misspecification in the first moment is put to the power three for the third moment equation. Thus, the results might be very sensitive to the specification and not very informative of the actual skewness of net gross margin. An interesting extension of the current paper would be to apply the methodology of climate risk exposure estimation exposed in chapter 2 in order to investigate the impact of AES on downside climate risk exposure. The prerequisite for such an approach is to have access to farm GPS coordinates. Unfortunately, this was not possible in Ireland at the time of the writing because of the strict rules on data confidentiality.

5.9 References

Antle, John M (1983), 'Testing the stochastic structure of production: a flexible moment-based approach', *Journal of Business & Economic Statistics,* 1 (3), 192-201.

--- (1987), 'Econometric estimation of producers' risk attitudes', *American Journal of Agricultural Economics,* 69 (3), 509-22.

--- (1989), 'Nonstructural risk attitude estimation', *American Journal of Agricultural Economics,* 71 (3), 774-84.

Antle, john M. (2010), 'Asymmetry, Partial Moments, and Production Risk', *American Journal of Agricultural Economics,* 92 (5), 1294-309.

Bangwayo-Skeete, Prosper F, Bezabih, Mintewab, and Zikhali, Precious (2012), 'Crop biodiversity, productivity and production risk: Panel data micro-evidence from Ethiopia', *Natural Resources Forum* (36: Wiley Online Library), 263-73.

Bar-Shira, Z., Just, R. E., and Zilberman, D. (1997), 'Estimation of farmers' risk attitude: an econometric approach', *Agricultural Economics,* 17 (2–3), 211-22.

Berentsen, PBM, Kovacs, K, and Van Asseldonk, MAPM (2012), 'Comparing risk in conventional and organic dairy farming in the Netherlands: an empirical analysis', *Journal of dairy science,* 95 (7), 3803-11.

Berry, Charles H. (1971), 'Corporate Growth and Diversification', *Journal of Law and Economics,* 14 (2), 371-83.

Centre, Rural Economy Research (2009), 'National Farm Survey, Definition of Variables for Data Output System', in Teagasc (ed.), (Athenry).

Chavas, Jean-Paul and Holt, Matthew T. (1990), 'Acreage Decisions under Risk: The Case of Corn and Soybeans', *American Journal of Agricultural Economics,* 72 (3), 529-38.

Chavas, Jean-Paul and Holt, Matthew T (1996a), 'Economic behavior under uncertainty: A joint analysis of risk preferences and technology', *The review of economics and statistics*, 329-35.

Chavas, Jean-Paul and Holt, Matthew T. (1996b), 'Economic Behavior Under Uncertainty: A Joint Analysis of Risk Preferences and Technology', *The Review of Economics and Statistics,* 78 (2), 329-35.

DAFF (2005), 'Scheme and Service 2005', (Department of Agriculture, Food and Fisheries, Ireland).

Damianos, Dimitri and Giannakopoulos, Nicholas (2002), 'Farmers' participation in agri-environmental schemes in Greece', *British Food Journal,* 104 (3/4/5), 261-73.

Di Falco, Salvatore and Chavas, Jean-Paul (2006), 'Crop genetic diversity, farm productivity and the management of environmental risk in rainfed agriculture', *European review of agricultural economics,* 33 (3), 289-314.

--- (2009), 'On crop biodiversity, risk exposure, and food security in the highlands of Ethiopia', *American Journal of Agricultural Economics,* 91 (3), 599-611.

Dupraz, Pierre, et al. (2003), 'The environmental supply of farm households: a flexible willingness to accept model', *Environmental and resource economics,* 25 (2), 171-89.

Emerson, Harriet J. and Gillmor, Desmond A. (1999), 'The Rural Environment Protection Scheme of the Republic of Ireland', *Land Use Policy,* 16 (4), 235-45.

Finger, Robert (2014), 'Risk considerations in the economic assessment of low-input crop production techniques: an example from Swiss wheat production', *International Journal of Agricultural Resources, Governance and Ecology,* 10 (1), 63-77.

Foudi, Sébastien and Erdlenbruch, Katrin (2011), 'The role of irrigation in farmers' risk management strategies in France', *European Review of Agricultural Economics*.

Gardebroek, Cornelis (2006), 'Comparing risk attitudes of organic and non-organic farmers with a Bayesian random coefficient model', *European Review of Agricultural Economics,* 33 (4), 485-510.

Gardebroek, Cornelis, Chavez, María Daniela, and Lansink, Alfons Oude (2010), 'Analysing production technology and risk in organic and conventional Dutch arable farming using panel data', *Journal of Agricultural Economics,* 61 (1), 60-75.

Groom, Ben, et al. (2008), 'The story of the moment: risk averse cypriot farmers respond to drought management', *Applied Economics,* 40 (3), 315-26.

Hynes, Stephen and Garvey, Eoghan (2009), 'Modelling Farmers' Participation in an Agri-environmental Scheme using Panel Data: An Application to the Rural Environment Protection Scheme in Ireland', *Journal of Agricultural Economics,* 60 (3), 546-62.

Just, David R., Khantachavana, Sivalai V., and Just, Richard E. (2010), 'Empirical Challenges for Risk Preferences and Production', *Annual Review of Resource Economics,* 2 (1), 13-31.

Kim, Kwansoo and Chavas, Jean-Paul (2003), 'Technological change and risk management: an application to the economics of corn production', *Agricultural economics,* 29 (2), 125-42.

Koundouri, Phoebe, et al. (2009), 'The effects of EU agricultural policy changes on farmers' risk attitudes', *Eur Rev Agric Econ*, jbp003.

Kumbhakar, Subal and Tsionas, Efthymios (2010), 'Estimation of production risk and risk preference function: a nonparametric approach', *Annals of Operations Research,* 176 (1), 369-78.

Kumbhakar, Subal C. (2002), 'Specification and Estimation of Production Risk, Risk Preferences and Technical Efficiency', *American Journal of Agricultural Economics,* 84 (1), 8-22.

Kumbhakar, Subal C. and Tveterås, Ragnar (2003), 'Risk Preferences, Production Risk and Firm Heterogeneity*', *The Scandinavian Journal of Economics,* 105 (2), 275-93.

Lence, Sergio H. (2009), 'Joint Estimation of Risk Preferences and Technology: Flexible Utility or Futility?', *American Journal of Agricultural Economics,* 91 (3), 581-98.

Ligon, Ethan and Schechter, Laura (2003), 'Measuring Vulnerability*', *The Economic Journal,* 113 (486), C95-C102.

Love, H. Alan and Buccola, Steven T. (1991), 'Joint Risk Preference-Technology Estimation with a Primal System', *American Journal of Agricultural Economics,* 73 (3), 765-74.

Mazorra, Angel Paniagua (2001), 'Agri-environmental policy in Spain. The agenda of socio-political developments at the national, regional and local levels', *Journal of rural studies,* 17 (1), 81-97.

Menezes, Carmen, Geiss, Charles, and Tressler, John (1980), 'Increasing downside risk', *American Economic Review,* 70 (5), 921-32.

Morris, J, Mills, J, and Crawford, IM (2000), 'Promoting farmer uptake of agri-environment schemes: the Countryside Stewardship Arable Options Scheme', *Land Use Policy,* 17 (3), 241-54.

Murphy, Geraldine, et al. (2014), 'An investigation into the type of farmer who chose to participate in Rural Environment Protection Scheme (REPS) and the role of institutional change in influencing scheme effectiveness', *Land use policy,* 39, 199-210.

O'donoghue, Cathal and Howley, Peter (2012), 'The Single Farm Payment: A Basic Income for Farmers?', *Basic Income Studies,* 7 (1).

Peerlings, Jack and Polman, Nico (2009), 'Farm choice between agri-environmental contracts in the European Union', *Journal of Environmental Planning and Management,* 52 (5), 593-612.

Pope, Rulon D. and Just, Richard E. (1991), 'On Testing the Structure of Risk Preferences in Agricultural Supply Analysis', *American Journal of Agricultural Economics,* 73 (3), 743-48.

Pratt, John W (1964), 'Risk aversion in the small and in the large', *Econometrica: Journal of the Econometric Society*, 122-36.

Saha, Atanu (1997), (Oxford University Press).

Saha, Atanu, Shumway, C. Richard, and Talpaz, Hovav (1994), 'Joint Estimation of Risk Preference Structure and Technology Using Expo-Power Utility', *American Journal of Agricultural Economics,* 76 (2), 173-84.

Serra, Teresa, Zilberman, David, and Gil, José M (2008), 'Differential uncertainties and risk attitudes between conventional and organic producers: the case of Spanish arable crop farmers', *Agricultural Economics,* 39 (2), 219-29.

Simtowe, Franklin, et al. (2006), 'Can Risk-aversion Toward Fertilizer Explain Part of the Non-adoption Puzzle for Hybrid Maize? Empirical Evidence from Malawi', *Journal of Applied Sciences,* 6 (7), 1490-98.

Vanslembrouck, Isabel, Huylenbroeck, Guido, and Verbeke, Wim (2002), 'Determinants of the Willingness of Belgian Farmers to Participate in Agri-environmental Measures', *Journal of agricultural economics,* 53 (3), 489-511.

Vollenweider, Xavier, Di Falco, Salvatore, and O'Donoghue, Cathal (2011), 'Risk preferences and voluntary agri-environmental schemes: does risk aversion explain the uptake of the Rural Environment Protection Scheme?', *Grantham Research Institute on Climate Change and the Environment,* Working paper no 48.

Voors, Maarten J, et al. (2012), 'Violent conflict and behavior: a field experiment in Burundi', *The American Economic Review,* 102 (2), 941-64.

Wynn, Gerard, Crabtree, Bob, and Potts, Jacqueline (2001), 'Modelling farmer entry into the environmentally sensitive area schemes in Scotland', *Journal of agricultural economics,* 52 (1), 65-82.

5.10 Appendix

Table 5.8: Estimates of the net gross margin moments equations

	Mean	Variance	Skewnesse (rescaled by 1 mio)
capital	-0.0169	186.2	876077.3
	(0.0214)	(249.4)	(683770.4)
capital^2	-5.43e-08	-0.0000454	
	(8.10e-08)	(0.000911)	
labour	-9210.1**	-74228454.4	1.52365e+11
	(4438.9)	(52314472.3)	(1.10796e+11)
capital×labour	0.0310**	-596.6***	
	(0.0132)	(175.6)	
land	111.7	419530.0	-1.29425e+09
	(99.13)	(964758.7)	(1.46642e+09)
capital × land	0.0000835	8.303***	
	(0.000346)	(3.022)	
land × land	-0.330*	-1018.4	
	(0.180)	(2008.4)	
land × labour	62.65	62903.2	
	(54.78)	(333385.3)	
labour × labour	1708.2	31683436.1**	
	(1283.5)	(15303934.2)	
diversification	59.07**	19883.7	6924.8***
	(26.74)	(178561.8)	(1759.1)
REPS	-6789.1***	-41505350.7**	-2.59249e+10
	(1760.0)	(18093153.7)	(1.41397e+11)
cattle	-8100.8***	-78487590.9***	1.28648e+11
	(2207.3)	(25039336.6)	(1.82355e+11)
REPS × cattle	6379.7***	44070623.7**	8.82768e+10
	(1987.1)	(18234172.6)	(1.57853e+11)
Tillage	-6917.1*	-89706524.6***	4.60331e+11**
	(3609.3)	(33275915.4)	(2.14258e+11)
REPS × tillage	1296.6	4466190.4	-2.35216e+11
	(3347.0)	(30593272.8)	(3.02152e+11)
sheep	-7884.5***	-83493565.2***	1.66581e+11
	(2311.0)	(26067852.7)	(2.07045e+11)
REPS× sheep	7961.0***	64702617.5***	-5.50756e+10
	(2320.5)	(20222066.7)	(1.77489e+11)
2007	8121.8***	32457453.4***	-5.40840e+10
	(490.2)	(4707567.1)	(3.29716e+10)
2008	4723.9***	18803710.1***	2.17755e+10
	(476.4)	(4771971.3)	(3.11914e+10)
2009	-8441.8***	59624028.6***	3.99888e+11***
	(551.6)	(6366744.1)	(4.08223e+10)
Constant	51028.0***	117464769.3**	-6.11541e+11***
	(5344.6)	(57509185.4)	(2.12922e+11)
Observations	3973	3708	3575
R^2	0.301	0.073	0.078

Fixed effect estimator with standard errors robust to clustering at the farm level in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$