Risk Preferences and Environmental Uncertainty: Implications for Crop Diversification Decisions in Ethiopia*

Mintewab Bezabih[†] Mare Sarr[‡]

April 29, 2010

Abstract

To the extent that diversifying income portfolio is used as a strategy for shielding against production risk, both individual risk preferences and weather uncertainty could affect crop diversification decisions. This paper is concerned with empirically assessing the effects of risk and rainfall variability on farm level diversity. Unique panel data from Ethiopia consisting of experimentally generated risk preferences measures combined with rainfall data are employed in the analysis. The major contribution of this study is its explicit treatment of individual risk preferences in the decision to diversify, which, to our knowledge has not been controlled for in previous similar empirical analyses. Both covariate shocks from rainfall variability and individual risk aversion are found to positively contribute to an increased level of diversity. These results imply that in-situ biodiversity conservation could be effective in areas with high rainfall variability. However, reduction in risk aversion which is associated with poverty reduction, is likely to reduce in-situ conservation.

JEL Classification: Q57, Q56, C33, C35

Keywords: Crop diversity, Experimental risk preferences, Rainfall, Non linear-IV

1 Introduction

Risk exposure and risk management are inherent components of agricultural activities. Farmers face various forms of risks, ranging from natural environmental uncertainty such as vagarious climatic conditions (drought, flood, etc.), pests and pathogens, to market-related factors like price volatility. In the presence of efficient insurance markets, farmers may insure themselves effectively

*We would like to thank Salvatore di Falco for his helpful comments. The usual disclaimer applies.

†CEMARE, University of Portsmouth, St. George's Building 141 High Street, PO12HY, Portsmouth, United Kingdom. Mintewab.bezabih@port.ac.uk

‡School of Economics and Environmental Policy Research Unit, University of Cape Town, Private Bag, Rondebosch 7700, South Africa. Email: mare.sarr@uct.ac.za

to manage these risks. However, in the absence of perfect insurance markets as is often the case in developing countries, exposure to such risks is likely to affect the *ex ante* production choices (Fafchamps 1992; Chavas and Holt 1996; Kurosaki and Fafchamps 2002).

Thus, there is evidence that in developing countries, farmers' choices for crop diversification may reflect an insurance mechanism designed to reduce production risk. Indeed, a growing body of research suggests that agro-biodiversity contributes to increased agricultural crop yield, and to reduced production risk (Smale *et al.* 1998; di Falco and Chavas 2006, 2008, 2009; Bezabih *et al.* 2009). Exploring the link between crop genetic diversity and wheat productivity in Pakistan's Punjab, Smale et al (1998) find a positive correlation with the mean yield and a negative correlation with the variance of the yield. This finding has been confirmed among others by di Falco and Perrings (2005), and di Falco and Chavas (2006). Furthermore, using a moment-based approach to estimate the mean, variance and skewness of a stochastic crop production function, di Falco and Chavas (2009) maintain that agro-biodiversity enhances farm-level productivity and reduces the risk of crop failure ("downside risk exposure" measured by the skewness) in Tigray, even when diversity has a variance-enhancing effect.[1]

These findings come to bolster ecologists' earlier arguments that greater diversity is associated with increased biomass (Tilman and Downing 1994; Tilman, Wedin and Knops 1996) and productivity of ecosystems (Tilman, Polasky, and Lehman 2005). These findings may be explained by two different channels: (i) the "sampling effect" hypothesis posits that a collection with species of diverse productivity is more likely to include high productivity species; and (ii) the "niche differentiation effect" states that as a collection contains more diverse species, ecological niches are likely to be better utilized.

From a methodological point of view, the measures of diversity are potentially endogenous, since the share of land allocated to a specific crop is a choice variable. In controlling for potential endogeneity bias, Benin et al 2003 and Bezabih et al 2009 estimate not only the productivity-diversity relationship but also the determinants of agro-biodiversity. Using various estimation methods, Bezabih et al 2009 find rainfall level to hinder agro-biodiversity, in that farmers who benefit from greater precipitations tend to engage less in crop diversification.

Interestingly, recent studies have found that the effect of diversity on productivity depends on rainfall so that agro-biodiversity has a larger impact at lower precipitation levels (di Falco and Chavas 2008; Bezabih *et al.* 2009). This means that farmers' choices of conserving agro-biodiversity is even more crucial in a water-stressed environment. This result implies that diversity may constitute an even greater insurance mechanism for the farmers facing production risk (crop failure) due to the vagaries of the weather. This is crucial in a country such as Ethiopia which has suffered severe droughts over the past few decades.

In addition to rainfall as an important determinant of crop diversity, standard economic theory also emphasizes that the tendency to diversify income portfolio is driven by individual risk preferences of decision makers.[2] While reviewing the literature relating individual risk preferences on optimal

[1]The variance-enhancing effect of diversity is at odds with the findings cited above. However, di Falco and Chavas (2009) show that what really matters, as far as risk exposure is concerned, is to establish which, between the variance and the skewness, dominates.

[2]It should be noted that individual risk preferences could differ across decision makers even when the overall riskiness of the decision making environment is constant. In the case we are studying, farmers living in the same

portfolio choice is too broad and beyond the purpose of our task here, some of the few examples include Kapteyn *et al.* (2002) who show that the optimal portfolio of households in the context of incomplete portfolio are explained by measured individual risk preferences. In analyzing worker's share holdings, Blasi *et al.* (2008) argue that diversification of each worker's entire portfolio is most efficient when in line with individual risk preferences. In addition, Heaton and Lucas (2000) show that the presence of background risks dictates differences in portfolio holdings in terms of stocks versus safe securities.

In agriculture, Fafchamps (1992) argues that exposure to risk affects the *ex ante* production choices, and optimal crop allocation depends among others on farmers' attitudes toward risk. Despite the importance of risk exposure in the economic literature on agro-biodiversity, however, very limited exposition exists on how farmers' idiosyncratic risk preference affects their choice of diversity. This paper sets out to extend existing analyses on the determinants of crop diversity using risk preferences and rainfall patterns as key factors.

The analysis starts out by setting up a simple theoretical framework that generates a testable hypothesis regarding the role of individual risk preference and rainfall variability on crop diversification decisions. The empirical analysis employs a Poisson random effects, quasi-fixed effects and non-linear instrumental variable specifications to empirically assess these relationships. The underlying reason exploring alternative specifications is to examine the potential endogeneity of the risk with respect to diversity. With no endogeneity, a standard random effects Poisson estimator would be a preferable fit while endogeneity of the risk variable implies a non-linear instrumental variables specification (Quadratic Variance Function method), or a quasi-fixed effects specification if the source of endogeneity is restricted only to the unobserved heterogeneity. The data source employed in this analysis is the Sustainable Land Management Survey conducted in the year 2005 and 2007 in Zones in the Amhara National Regional State of Ethiopia, that consists of experimentally generated risk preferences measures combined rainfall data from the Ethiopian Meteorological Authority.

The major contribution of this study is its explicit treatment of individual risk preferences in the decision to diversify, which, to our knowledge has not been controlled for in previous similar empirical analyses. The results of the analysis demonstrated that both rainfall availability and risk preferences increase the level of diversified crop portfolio while rainfall variability has less significant impact on diversity.

The paper is structured as follows: section 2 discusses the background literature on risk and crop diversification. A simple theoretical model between diversity and its key determinants is presented in Section 3. Section 4 presents the empirical/econometric methodology while section 5 provides the data and statistical analysis. Results are reported and discussed in section 6. Finally, section 7 concludes.

2 Background to Risk and Diversity

Insurance markets are conspicuously missing in the agricultural sector in many developing countries (Fafchamps 1992; Dercon 1996). Since farmers are unable to transfer the risk they are facing

---

village, while exposed to the same rainfall patterns could still have differing individual risk preferences.

to a third party, they must find both risk-management and risk-coping strategies to hedge against potential adverse shocks. For instance, diversifying land allocation thereby increasing crop agro-biodiversity, or growing less risky crops are typical risk-management strategies. The extent to which these strategies are adopted may depend on farmers' risk aversion (Fafchamps 1992).

Farmers' motives for investing in agro-biodiversity have traditionally been explored within the so-called household models. Those models seek to establish the main determinants of crop variety choice (Fafchamps 1992) and agro-biodiversity using a utility maximizing farm household framework (Herath *et al.* 1982; Adesina and Zinna 1993; Barkley and Porter 1996; Van Dusen 2000; Smale *et al.* 2001; Benin *et al.* 2003). Households are both consumers (of finished goods or rental of labor) and producers—using their endowments, labor, land and capital to grow crops. The choice variables of interest are typically land allocation (shares of land area devoted to specific crops) or production levels. These decisions generally depend on a number of parameters—namely prices (including wages), income levels, and various socio-economic, physical and geographical characteristics—which in turn affect the level of agro-biodiversity as diversity indices are often constructed from the area shares (Van Dusen 2000; Benin *et al.* 2003). This allows applied economists to estimate agro-diversity as a function of these choice variables. For example, using the Margalef index of richness (number of species or varieties) or the Shannon index of evenness as a measure of diversity, Benin *et al.* (2003) find that households with more male labor, more oxen or larger farms grow more diverse cereal crops, while Bezabih *et al.* (2009) suggest that the level of rainfall and household endowments tend to govern crop diversity decisions rather than plot or other household characteristics.

However, these models, surprisingly, do not generally include farmers' risk preference. In household models, Fafchamps (1992) is certainly one of the rare studies that brings farmers' risk attitudes to the fore although the study is not particularly concerned with agro-biodiversity per se. Investigating why large-scale farmers in the developing world tend to allocate larger shares to cash crops than small-scale farmers, Fafchamps highlights an interesting relationship between farmers' crop portfolio diversification, idiosyncratic risk aversion and consumption preferences. He develops a dynamic model in which he shows that the effect of risk aversion on crop portfolio is a priori ambiguous and results from the combination of: (i) a direct portfolio effect; and (ii) a consumption effect by which a risk averse farmer will insure himself against potential price risk by weighing his crop allocation in favor of the crops with a high consumption price. The overall direction of the effect will hinge upon the parameters of the model.

More recently, Quaas and Baumgärtner (2008) touch on the very issue of the link between agro-biodiversity and risk preference at a rather abstract level. Their main motivation is to understand how farmers manage their agro-biodiversity portfolio to insure themselves against income risk, and how the availability of financial insurance may impact their management decision. They develop an interesting ecological-economic model of a farmer's choice of biodiversity in the presence of a financial insurance sector—unlike most of the literature dealing with missing insurance markets in developing countries. They show among others that agro-biodiversity is used by farmers as a natural insurance mechanism (as opposed to financial insurance) and increases with risk aversion in an unambiguous manner. The choice variable in this paper differs from the earlier contributions cited above in that the farmer does not choose the crop allocation which in turn is associated with agro-biodiversity. Instead, in Quaas and Baumgärtner (2008), the farmer chooses the level of biodiversity, which is clearly a more abstract control variable. In our paper, we combine insights from both Quaas and Baumgärtner's modeling and the crop choice household models to suggest a simple model that will be used to analyse the relationship between crop diversity and farmers' risk preferences.

3 Simple model

Assume a farmer grows $i \in [1,n]$ different crops in a risky environment (e.g. weather, diseases or pests) so that crop yield q_i is a random variable. The farmer may impact the distribution of the crop yield and, thus on his income, by choosing the area of land L_i allocated to each crop i, with $\sum_i L_i = \bar{L}$ and $\bar{L}$ being the total land area. Let X represent the other forms of inputs required to grow crops. It is also assumed that the mean $\mu_i(L_i,X)$ and variance $\sigma_i(L_i,X)$ of the crop yield depend on land allocation L_i and X such that:

$$\mathbb{E}(q_i) = \mu_i(L_i,X), \ \frac{\partial \mu_i}{\partial L_i}(L_i,X) > 0, \ \frac{\partial^2 \mu_i}{\partial L_i^2}(L_i,X) < 0 \tag{1}$$

$$Var(q_i) = \sigma_i^2(L_i,X), \ \frac{\partial \sigma_i}{\partial L_i}(L_i,X) \lesseqgtr 0, \ \frac{\partial^2 \sigma_i}{\partial L_i^2}(L_i,X) > 0 \tag{2}$$

These assumptions indicate that the mean of the yield increases in land allocation and is concave, while the variance may either increase or decrease with land allocation and is convex. Assume also that the direct and opportunity costs of production $C_i(L_i,X)$ are increasing and convex.

The farmer's net income is therefore given by $y = Q - C(L,X)$, where $Q = \sum_i q_i(L_i,X)$ and $C(L,X) = \sum_i C_i(L_i,X)$. The farmers are non-satiated and risk-averse with respect to income. Their preferences are represented by a von Neumann-Morgenstern expected utility function

$$U = \mathbb{E}[u(y)] \tag{3}$$

where $u(y)$ is a Bernoulli utility function which is assumed to be increasing ($u' > 0$) and strictly concave ($u'' < 0$). Maximising the expected utility is equivalent to maximising the certainty equivalent $CE = \mathbb{E}(y) - R$ (see Chavas 2004):

$$\begin{cases} \max_{L_i,X} & CE = \mathbb{E}(y) - R \\ \text{s.t.} & y = Q - C(L,X) \\ & Q = \sum_i q_i(L_i,X) \\ & \sum_i L_i = \bar{L} \end{cases} \tag{4}$$

where $R \approx \frac{A}{2} Var(y)$ is the risk premium and A is the Arrow-Pratt measure of absolute risk aversion. The certainty equivalent can be re-written as:

$$CE = \sum_i [\mu_i(L_i,X) - C_i(L_i,X)] - \frac{A}{2}\left[\sum_i \sigma_i^2(L_i,X) + 2\sum_{i<j}\rho\sigma_i(L_i,X)\sigma_j(L_j,X)\right]$$

where $\mathbb{E}(y) = \sum_i [\mu_i(L_i,X) - C_i(L_i,X)]$ and

$$Var(y) = Var\left(\sum_i q_i(L_i,X)\right) = \sum_i Var_i(L_i,X) + 2\sum_{i<j} Cov(q_i,q_j) = \sum_i \sigma_i^2(L_i,X) + 2\sum_{i<j}\rho\sigma_i(L_i,X)\sigma_j(L_j,X).$$

The farmer then chooses the level land allocation and input levels that maximise the following problem:

$$\max_{L_i,X} \sum_i [\mu_i(L_i,X) - C_i(L_i,X)] - \frac{A}{2}\left[\sum_i \sigma_i^2(L_i,X) + 2\sum_{i<j}\rho\sigma_i(L_i,X)\sigma_j(L_j,X)\right] \tag{5}$$

The first order condition with respect to L_i gives:[3]

$$\frac{\partial\mu_i}{\partial L_i}(L_i^\star,X) - \frac{\partial C_i}{\partial L_i}(L_i^\star,X) - A\frac{\partial\sigma_i}{\partial L_i}(L_i^\star,X)\left[\sigma_i(L_i^\star,X) + \rho\sigma_j(L_j,X)\right] = 0 \tag{6}$$

Given $L_i^\star$, the optimal share of land allocated to crop i is given by $l_i^\star = l_i^\star(A,\bar{L}) = L_i^\star/\bar{L}$ and depends on the farmer's risk preference. This decision will in turn affect crop agro-biodiversity as diversity indices are constructed from these area shares (Benin *et al* 2003).

$$D = D(l_i^\star(A,\bar{L})) \tag{7}$$

4 Empirical Methodology

This paper will investigate the relationship between agro-biodiversity and risk preference using evidence from Ethiopia. Like in many low income countries, agriculture occupies a large share of the Ethiopian economy with 40% of the GDP, 90% of exports, and 85% of employment (Bezabih et al. 2009). Ethiopian agriculture relies almost exclusively on rainfall. During the last thirty years, droughts, incidence of pests, as well as animal and human disease have been recurrent events (Dercon 2004). These events have contributed to increase the vulnerability of rural households and impede food security. In extreme cases, they have resulted in large scale famine as happened in the 1980s.

In such an uncertain environment, a staggering proportion of farmers seem to exhibit severe or extreme risk aversion, with poorer farmers being more averse to risk (Yesuf and Bluffstone 2009). In contrast, earlier studies in India and Zambia found that farmers were only moderately risk averse (Binswanger 1980, and Wik and Holden 1998). Such extreme risk attitude is likely to impact upon farmers' decisions to invest and land allocation (and therefore crop biodiversity) to increase productivity and insure against the risk of crop failure (Di Falco and Chavas 2009).

[3] The first order conditions with respect to X are not of interest so we do not compute them.

In this section we model the decision of a farming household to diversify its crop allocation using a limited dependent variable framework. Our dependent variable, y_{it}, represents household i's farm-level diversity in time t, where $i = 1,...,n$ and $t = 1,..T$. It is measured in terms of count diversity, i.e. the number of crops grown per farm.[4] Equation (8) lays out the relationships between the factors affecting the household's decision to diversify and the level of diversity at a farm level:

$$y_{it} = \beta \mathbf{x_{it}} + \gamma \mathbf{q_{it}} + \varphi r_{it} + \xi_{it} \tag{8}$$

where $\mathbf{x_{it}}$ denotes a vector of observable household socioeconomic characteristics and physical farm characteristics, $\mathbf{q_{it}}$ represent a vector of rainfall variables, and r_{it} represents household risk preference, a possibly endogenous variable. The coefficients β, γ and φ represent the respective vectors of parameter estimates and ξ_{it} represents the error term.

Estimation of equation (8) needs to take into account two issues: (i) the possible existence of unobserved farm-level heterogeneity; and (ii) the endogeneity of the risk preference variable. Rewriting equation (8) gives:

$$y_{it} = \beta \mathbf{x_{it}} + \gamma \mathbf{q_{it}} + \varphi r_{it} + \alpha_i + u_{it} \tag{9}$$

where the composite error term $\xi_{it} = \alpha_i + u_{it}$ is composed of a normally distributed random error term $u_{it} \sim \mathscr{N}\left(0, \sigma_u^2\right)$ and an unobserved household specific effects α_i.

Given that the observable covariates in equation (9) do not account for all the systematic variation in y_{it}, two alternative approaches have been suggested in the literature. The fixed effects estimator takes α_i to be a group specific constant term and uses a transformation to remove this effect prior to estimation. In addition, the fixed effects estimator allows for arbitrary correlation between α_i and the explanatory variables in any time period, i.e. $\mathbb{E}(\alpha_i|\mathbf{x_{it}}, \mathbf{q_{it}}, r_{it}) \neq 0$ (Wooldrige 2001).

The alternative approach of the random effects estimator has a similar specification to the fixed effects estimator. However, the random effects estimator does not rely on the data transformation in fixed effects estimator and the associated shortcoming of removing any time-invariant explanatory variables along with α_i (Wooldrige 2001). Conversely, the random effects model requires the regressors to be uncorrelated with the individual effects α_i, i.e., $\mathbb{E}(\alpha_i|\mathbf{x_{it}}, \mathbf{q_{it}}, r_{it}) = 0$. So long as this assumption is satisfied, the random effects estimator will be a consistent and efficient estimator (Baltagi 2001; Mundlak 1978).

A violation of the exogeneity assumption underlying the random effects model leads to biased parameter estimates. To remedy this, Mundlak (1978) suggests modeling explicitly the relationship between time varying regressors $\mathbf{z_{it}} = (\mathbf{x_{it}}, \mathbf{q_{it}}, r_{it})$ and the unobservable effect α_i in an auxillary regression. In particular α_{it} can be approximated by a linear function:[5]

$$\alpha_{it} = \omega \mathbf{s_{it}} + \gamma \mathbf{z_{it}} \tag{10}$$

[4] There are a number of diversity measures used in empirical agricultural studies. These include : (i) the Count index which measures species richness on a given farm; (ii) the Malgref Index which measures richness but accounts for land area; (iii) the Shannon Index which measures richness and relative abundance; and (iv) the Berger Parker Index which measures relative abundance (Benin *et al.* 2003). The Count index is the most commonly used index.

[5] Similar approaches are proposed by Hausman and Taylor (1981) and Baltagi *et al.* (2001)

where $\mathbf{s_{it}}$ represents a vector of explanatory variables and ω is a vector of parameters to be estimated. Averaging over t for a given i yields $\alpha_i = \omega \bar{s}_i + \gamma \bar{z}_i$, so that substituting the resulting expression into (9) gives:

$$y_{it} = \beta \mathbf{x_{it}} + \gamma \mathbf{q_{it}} + \varphi r_{it} + \omega \bar{s}_i + \gamma \bar{z}_i + u_{it} \tag{11}$$

Controling for these sources of unobserved heterogeneity by adding the means of time varying observed covariates is commonly known as the pseudo-fixed effects or the Mundlak-Chamberlain's Random Effects Model.

We start our analysis with the strong assumption of exogenous risk preference in the diversity equation. Given the discrete nature of y_{it}, we opt for a random effects Poisson model specification, which is akin to estimating equation (9) using random effects model.

When we allow for arbitrary correlation between the covariates and the household-specific effects, we then estimate equation (11). While the pseudo-fixed effects estimation removes the possible source of endogeneity associated with unobserved heterogeneity, addressing endogeneity caused by correlation with the dependent variable needs the instrumental variables estimation. The possible endogeneity of the risk preference variable calls for estimating equation (8) using the appropriate instrumental variable approach.[6] This implies specifying a relationship between the risk variable and a potential instrument.

$$r_{it} = \mu h_{it} + \theta m_{it} + v_{it} \tag{12}$$

where (μ, θ) is the vector of coefficients to be estimated, and h_{it} and m_{it} are the determinants of risk preference and v_{it} is the error term. Since diversity is a discrete variable, equation (8) is nonlinear in the instruments. A consistent estimator can be found using the Generalized Method of Moments (Davidson and MacKinnon 1993). A special case of a GMM estimator is a Generalized Linear Moments estimator (GLM)—originally developed by Nelder and Wedderburn (1972), and McCullagh and Nelder (1989)—which is a unified regression methodology for a variety of discrete, continuous and censored responses that are assumed to be independent. The quasi-likelihood variance function estimator (QVF) is a variant of the GLM estimator where the variance function is specified instead of the moments of the distribution (McCullagh and Nelder 1989; Carroll, Ruppert and Stefanski 1995).[7] Examples of empirical studies that use QVF include Hagedoorn and Hesen (2009) and Wagstaff and Lindelow (2008).

Testing the validity of instruments is undertaken using the Amemiya–Lee–Newey (ALN) overidentification test and the Hausman-Hu test for endogeneity. Unlike the Generalized Linear Model, the QVF command in Stata does not support the predict command and other post estimation methods. Alternative approaches include 2SLS estimation and testing various aspects of instrumental variable regressions such as the relative weakness of the instruments (Hagedoorn and Hesen 2009).

[6] To the extent that the sources of endogeneity are unobserved fixed effects, there would not be a need to carry out the instrumental variable estimation.

[7] In terms of estimation procedure, the QVF model is similar to a generalized linear model (GLM in Stata) but has the ability to include instrumental variables, a functionality that was added to address measurement error, but may be utilized by the user for other purposes (Hardin, Schmiediche and Carroll 2003) such as correcting for endogeneity (Hagedoorn and Hesen 2009).

5 Data

The data source employed in this analysis is the Sustainable Land Management Survey conducted in the year 2005 and 2007 in Zones in the Amhara National Regional State of Ethiopia. A total of 12 villages were included in the study, six from the East Gojjam Zone and the other six from the South Wollo Zone. The East Gojjam Zone is a high agricultural potential region with differing agro-ecology and slightly differing cropping patterns from the South Wollo Zone. The dataset contains information on household socio-economic characteristics, physical farm characteristics, production, and risk preference. The description of the variables as well as the summary statistics of the data used to estimate the diversity and risk preference equations are provided in Tables 1 and 2.

Our dependent variable, farm level diversity, is defined as the number of crops grown by the household. This diversity index, also known as count index, would be one for a household that grows only one type of crop and could be greater than unity for households that have more than one type of crop in their farm. Information on the level of farm diversity shows that there has been a moderate change in the level of diversity over time.

A hypothetical risk preference experiment was conducted with all the respondents, paired with the main survey using a lottery choice experiment where respondents were presented with a choice of six pairs of farming systems. Each choice consists of a pair of good and bad outcomes, each outcome occurring with a probability of 50%. This enables the calculation of the expected gains (i.e., the average of the two outcomes), and the spread (i.e., the difference between the two outcomes). The categories in the risk preference experiment represent the extent to which respondents are willing to take up risky choices. Accordingly the extreme risk aversion category represents households who are willing to take the smallest spread in gains and losses, followed by severe, moderate, intermediate, and slight risk aversion categories, while the neutral risk aversion category corresponds to respondents willing to take the biggest spread in gains and losses. The choice sets were arranged in a tree structure, where the choice made in the first choice set determined whether one branches into a more or less risky choice alternative. For instance, choosing a moderately risky alternative in a choice set followed by the choice of a less risky alternative would indicate the individual's risk preference as moderate. Table 3 shows the complete choice sets presented to the respondents.[8]

The risk index is a categorical variable constructed by assigning a value $j \in \{1,2,...,6\}$ to each of the six categories of risk preference as shown in the last column of Table 3. Higher values of the risk index indicate to higher levels of risk aversion. Typically, the risk preference index equals one for "neutral risk preference" (i.e., the lowest risk aversion category) and gradually increases to six for the "extreme risk aversion" category.

Table 2 summarizes the descriptive statistics for the variables used in the regression analyses. The average risk preference index is 4.08 in 2005 and 3.88 in 2007, indicating that the average farmer is moderately risk averse. The total annual rainfall has increased by almost 50% from 1041 mm in 2005 and 1504 mm in 2007. Rainfall variability, which is computed as the ratio of the mean to

[8] A bad harvest ranged from between a 0 kg output to a 100 kg output, while a good harvest ranged between 100 and 400 kg. An extreme outcome consisted of an expected gain of 100 kg and a spread of 0 kg, while a neutral outcome consisted of an expected gain of 200 kg and a spread of 400 kg.

the variance of monthly rainfall for a given year, has remained fairly constant around 1.20 mm. It provides an additional information about the risk due to rainfall.

Numerous household characteristics are controled for in our analysis. The typical average household in 2005 is fairly large and composed of 3 adult males and 3 adult females, suggesting that there is adequate supply of both male and female labor. Households are typically headed by males aged somewhat around 50 with a low education level—only a third of them are able to read and write. Female headed households represent less than 20% of the sample. Regarding wealth measures, the sampled households own on average between 2 and 3 oxen, and 4 tropical livestock units.

The physical characteristics of the farms provide us with further key information that we want to control for. The average size of holdings is about 1.3 hectare and contains on average 2.4 fertile plots and 0.73 flat slopped plots with little variation between 2005 and 2007, suggesting a relatively homogeneous distribution of fertile land.

Since the original data are collected at plot levels, and since our primary interest is in farm level diversity, we have computed annual variables focusing on the socioeconomic characteristics of the household head.[9]

6 Results

The empirical analysis investigates the relationship between farm level diversity, risk preferences and rainfall patterns. We estimate this relationship using the three models discussed in section 4: 1) the standard random effects Poisson model; 2) the pseudo-fixed effects Poisson which controls for possible unobserved heterogeneity at household/farm level; and 3) the quadratic variance function (QVF) which controls for the possible endogeneity of the risk preference measure using an instrumental variable approach. The coefficient estimates and associated standard errors for the three models are presented in Table 4.

We first discuss the results of the standard random effects Poisson model presented in Columns 2 and 3. We find that total annual rainfall is a positive and significant determinant of diversity. While this finding seems counter-intuitive, it can be explained by the severely rainfall constrained nature of Ethiopian agriculture. Given a reasonable degree of moisture (rainfall) abundance, increased average rainfall could result in reduced diversity as farmers would focus on crops that do particularly well with abundant rainfall, compared to the reference point. However, since rainfall availability is generally a constraint in Ethiopia, increased availability of rainfall could be seen as an opportunity to increase the range of crops grown. Again, due to our reference point of moisture constraint, the opposite effect is likely to hold for drier conditions. Reduced average rainfall is likely to limit households to growing only 'safer' crops that do well in moisture constrained conditions, leading to reduced level of diversity. This result supports the findings by Di Falco and Chavas (2008) that the interaction between diversity and rainfall lead to increased productivity.

We also find that increased rainfall variability significantly increases the decision to diversify. , Households facing increased increased weather uncertainty may choose to increase the spectrum

[9] Such approaches are common in agricultural household studies.

of crops—both those that fare well under drier conditions and those that fare well under wetter conditions.

In sum, both increased average rainfall and its wider spread are likely to increase diversity. By implication the effect of reduced average rainfall and its narrower range is to reduce the level of diversity.

In addition to these sources of covariate risks, we account for individual risk preferences captured by the risk index constructed using the experimental risk preference measures. We find that risk aversion is positively associated with diversity—albeit weakly, i.e., at 13% level of significance. This implies that more risk averse farmers tend to favor greater crop diversity. Thus, greater individual risk aversion and greater external risks (captured here by greater rainfall availability and variability) lead to greater crop diversity.

Of the socio-economic characteristics, age and number of adult males and females in the household are not significantly associated with the level of diversity. Education, measured as the ability of the head of the household to read and write, is also insignificant. However, measures of wealth such as oxen and livestock ownership are significant determinants of diversity.[10] Benin *et al.* (2003) note, with respect to low income agriculture, that wealth might positively affect the decision to diversify. This is because greater wealth tends to make farmers more willing to use opportunities to diversify which in turn increases their income.

Regarding farms physical characteristics, farmers with larger farms tend to select more diverse crops. This finding is in line with Benin *et al.* (2003). In addition, while farms with more fertile plots have a positive and statistically significant relationship with diversity, farms with flat slopes have a negative although statistically insignificant association. To control for possible non linearities in the socio-economic and farms physical characteristics, square terms of adult male and female labor, oxen, livestock and farm size are included as additional explanatory variables. Only the square term of farm size is significant, suggesting that diversity increases with land area at a decreasing rate.

It should be noted that while village level and agro-ecological level variables may have a significant impact on the decision to diversify, we are not able to directly control for them because of their correlation with the rainfall variables.

Columns 3 and 4 of Table 4 present alternative estimates of the diversity regressions using the pseudo-fixed effects model. We include the mean of the following time varying household covariates: oxen, livestock, head's age, adult male, adult female, and farm size. We find similar results both in terms of the direction of the relationships and the magnitudes. Furthermore, the mean oxen and the mean farm size are statistically significant, indicating the presence of unobserved fixed effects.

In Columns 5 and 6 of Table 4, we estimate a QVF model to account for the possible endogeneity of farmers' individual risk preferences. To this effect, an instrumental variables approach is used. We use tenure (in)security, as our instrument for the risk index. It is measured as the expectation of changes in the size of the holdings (increase or decrease) in the coming five years.

The results are qualitatively very similar to those in the previous specifications. Controling for endogeneity has somewhat improved our results as the coefficient of the risk index has now gained

[10] In most agricultural studies on Ethiopia, it is common to use livestock as a proxy for wealth.

statistical significance at the 5% level. The main message is that both rainfall levels, rainfall variability, and risk aversion enhance crop diversity. This suggests that in the absence of insurance markets, farmers tend to diversify their crop portfolio as These findings suggest that in the absence of insurance markets, farmers select a diverse crop portfolio as an insurance mechanism to cope with both increased idiosyncratic risk and aggregate risk

To ensure the soundness of our instrumental variables approach, we run an exogeneity test and an over-identification test. The results suggest that the conventional statistical tests for sound instruments are satisfied. Indeed, the Hausman-Hu exogeneity test rejects the null hypothesis that the risk index can be treated as exogenous (p-value=0.00). Besides, the Amemiya-Lee-Newey over-identification test (Baum *et al.* 2006) indicates that the null hypothesis that the instruments are uncorrelated with the error term and correctly excluded from the outcome equation, cannot be rejected. The QVF model is therefore our preferred specification.

7 Conclusion

This study analyzes the links between farm-level crop diversification decisions and risk factors using farm household data from the central highlands of Ethiopia and experimental measures of risk preferences merged with rainfall data. Previous empirical analyses of the decision to diversify have focused on household and farm characteristics, and, to some extent, on rainfall variability. However, to our knowledge, no other study has looked into the impact of individual risk preferences in crop diversification decisions using measures of risk preferences. Accordingly, the major objective is to address whether diversity is responsive to individual households' risk preferences or whether it is more responsive to rainfall patterns, a proxy for covariate risks.

This study develops a framework for assessing the impact of households' attitude toward risk and rainfall on the decision to diversify. The framework is tested with alternative econometric specifications. The results of the analysis suggest that both rainfall availability, rainfall variability and risk preferences increase the level of diversified crop portfolio. The findings seem to be robust to a number of specifications and to the endogeneity of households' risk preference.

Empirical studies on farm level diversification decisions have also focused largely on crop diversity. Extending the risk-diversification interactions to all farm enterprises including livestock production and off-farm employment may further illuminate our understanding of the role of covariate and idiosyncratic risk in enterprise choice.

References

Aguirre Gómez, J.A., Bellon, M.R. and Smale, M. (2000): A regional analysis of maize biological diversity in Southeastern Guanajuato, Mexico. *Economic Botany*, 54(1): 60–72.

Adesina, A. and Zinnah, M. (1993): Technology Characteristics, Farmer Perceptions and Adoption Decisions: A Tobit Model Application in Sierra Leone. *Agricultural Economics*, 9: 297–311.

Baltagi, B.H. (2001). Econometric Analysis of Panel Data (Second Edition), New York, Wiley.

Barkley, A. and Porter, L. (1996): The Determinants of Wheat Variety Selection in Kansas, 1974 to 1993. *American Journal of Agricultural Economics*, 78: 202–211.

Baum, C. F., Schaffer, M. E., Stillman, S. and Wiggins, V. (2006): 'Overid: Stata module to calculate tests of overidentifying restrictions after ivreg,ivreg2, ivprobit, reg3', http://ideas.repec.org/c/boc/bocode/s396

Baumgärtner, S. (2007): The insurance value of biodiversity in the provision of ecosystem services. *Natural Resource Modeling*, 10(1): 87–127.

Baumgärtner, S. and Quaas, M. (2008): Agro-biodiversity as natural insurance and the development of financial insurance markets pp. 293–317, in Eds. Kontoleon A., Pascual, U. and Smale, M. , *Agrobiodiversity, Conservation and Economic Development*, Routledge.

Bellon, M.R. and Taylor, J.E. (1993): Folk soil taxonomy and the partial adoption of new seed varieties. *Economic Development and Cultural Change*, 41: 763–786.

Benin, S., Smale, M., Gebremedhin, B., Pender, J. and Ehui, S. (2003): Determinants of cereal crop diversity on farms in the Ethiopian Highlands. *Environment and Production Technology Division Discussion Paper 105*. International Food Policy Research Insititute (IFPRI), Washnigton DC.

Bezabih, M., di Falco, S. and Yesuf, M. (2009): Seeds for livelihood: Crop biodiversity and food production in Ethiopia." *Mimeo*.

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

Blasi J. R. , D. L. Kruse, H. M. Markowitz. 2008. Risk and Lack of Diversification under Employee Ownership and Shared Capitalism. *NBER Working Paper* No. 14229

Brush, S., Taylor, J.E. and Bellon, M.R. (1992): Biological diversity and technology adoption in Andean potato agriculture. *Journal of Development Economics*, 39: 365–387.

Carroll, R. J., Ruppert, D. and Stefanski, L. A. (1995): *Measurement Error in Nonlinear Models*. Chapman & Hall.

Chavas, J.-P. (2004): *Risk Analysis in Theory and Practice*. Academic Press, Elsevier.

Chavas, J.P., and Holt, M.T. (1996): Economic behavior under uncertainty: A joint analysis of risk preferences and technology. *Review of Economics and Statistics*, 78: 329–335.

Davidson, R. and MacKinnon, J.G. (1993): *Estimation and Inference in Econometrics*. Oxford University Press, New York, NY.

Dercon, S. (1996): Risk, crop choice, and savings: Evidence from Tanzania. *Economic Development and Cultural Change*, 44(3): 485–513.

Dercon, S. (2004): Growth and shocks: evidence from rural Ethiopia. *Journal of Development Economics*, 74(2): 309–329.

Di Falco, S. and Perrings, C. (2005): Crop biodiversity, risk management and the implications of agricultural assistance. *Ecological Economics*, 55: 459–466.

Di Falco, S., and Chavas, J-P. (2008): Rainfall shocks, resilience and the dynamic effects of crop biodiversity on the productivity of the agroecosystem. *Land Economics*, 84(1): 83-96.

Di Falco, S. and Chavas, J-P. (2009): On crop biodiversity, risk exposure and food security in the highlands of Ethiopia. *American Journal of Agricultural Economics*, 91(3): 599–611.

Eswaran, M., and Kotwal, A. (1990): Implications of credit constraints for risk behaviour in less developed economies. *Oxford Economic Papers*, 42(2): 473–482.

Fafchamps, M. (1992): Cash crop production, food price volatility, and rural market integration in the Third World. *American Journal of Agricultural Economics*, 74(1): 90–99

Hagedoorn, J. and Hesen, G. (2009): Contractual complexity of R&D alliances: A two-dimensional analysis of the determinants of contractual complexity. *SSRN Working Papers*. http:/ssrn.com/abstract=1367530.

Hardin, J., Schmiediche, H. and Carroll, R. (2003): Instrumental variables, bootstrapping, and generalized linear models. *The Stata Journal*, 3(4): 351–360.

Hausman, J.A. and Taylor, W.E. (1981): Panel data and unobservable individual effects. *Econometrica*, 49(6): 1377–1398.

Heaton, J. and Lucas, D. (2000): Portfolio choice in the presence of background risk. *Economic Journal*, 110 (1): 1–26.

Herath, H., Hardaker, J.B. and Anderson, J.R. (1982): Choice of varieties by Sri Lanka rice farmers: comparing alternative decision models. *American Journal of Agricultural Economics*, 64(1): 87–93.

Kaptyen, R. And F. Teppa. 2002. Subjective measures of risk aversion and optimal portfolio choice. Rand Labour and Population Program. *Working Paper Series*, 02-03 Tilburg University.

Kurosaki, T. and Fafchamps, M. (2002): Insurance market efficiency and crop choices in Pakistan. *Journal of Development Economics*, 67(2): 419–453

McCullagh, P. and Nelder, J.A. (1989): *Generalized Linear Models*. Chapman and Hall: London.

Mundlak, Y. (1978): On the pooling of time series and cross-section data. *Econometrica*, 46: 69–85.

Naima S. (2009): *Estimation of QVF measurement error models using empirical likelihood method. A Dissertation*. Graduate College of Bowling Green State University, USA.

Nelder, J. A. and Wedderburn, R.W.M. (1972): Generalized linear models. *J. Roy. Statist. Soc.* Ser. A 135: 370–384.

Revelli, F. (2010): Spend more, get more? An inquiry into English local government performance. *Oxford Economic Papers*, 62: 185–207

Smale, M., Hartell, J., Heisey, P.W. and Senauer, B. (1998): The contribution of genetic resources and diversity to wheat production in the Punjab of Pakistan. *American Journal of Agricultural Economics*, 80: 482–493.

Smale, M., Bellon, M. and Aguirre, A. (2001): Maize diversity, variety attributes, and farmers' choices in Southeastern Guanajuato, Mexico. *Economic Development and Cultural Change*, 50(1): 201–225.

Smale, M., Meng, E., Brennan, J.P. and Hu, R. (2002): Determinants of spatial diversity in modern wheat: Examples from Australia and China. *Agricultural Economics*, 28(1):13–26.

Tilman, D. and Downing, J.A. (1994): Biodiversity and stability in grasslands. *Nature*, 367: 363–365.

Tilman, D., Wedin, D. and Knops, J. (1996): Productivity and sustainability influenced by biodiversity in grassland ecosystems. *Nature*, 379: 718–720

Tilman, D., Polasky, S. and Lehman, C. (2005): Diversity, productivity and temporal stability in the economies of humans and nature. *Journal of Environmental Economics and Management*, 49(3): 405–426.

Van Dusen, E. (2000). *In Situ Conservation of Crop Genetic Resources in the Mexican Milpa System*. PhD Dissertation, University of California, Davis.

Wik, M., and Holden, S. (1998): Experimental studies of peasant's attitudes toward risk in Northern Zambia. *Discussion Paper D-14*, Department of Economics and Social Sciences, Agricultural University of Norway.

Wooldridge, J.M. (2002): *Econometric Analysis of cross section and panel data*. Cambridge, MA, USA: Massachusetts Institute of Technology.

Williams, R. (2006): Generalized ordered logit/partial proportional-odds models for ordinal dependent variables. *Stata Journal*, 6: 58–82.

Wagstaff, A. and Lindelow, M. (2008): Can insurance increase financial risk? The curious case of health insurance in China. *Journal of Health Economics*, 27 (2008) 990–1005.

Yesuf, M., and Bluffstone, R. (2009): Poverty, Risk aversion and path dependence in low income countries: Experimental evidence from Ethiopia. *American Journal of Agricultural Economics*, 91(4):1022–1037.

Table 1: Description of variables

Variables	Description
Dependent variable	
Diversity	Count Diversity (number of crops grown per farm)
Risk preference variables	
Neutral risk aversion	Household classified as neutral in risk aversion (dummy)
Slight risk aversion	Household classified as slightly risk averse (dummy)
Intermediate risk aversion	Household classified as intermediately risk averse (dummy)
Moderate risk aversion	Household classified as moderately risk averse (dummy)
Severe risk aversion	Household classified as severely risk averse (dummy)
Extreme risk aversion	Household classified as extremely risk averse (dummy)
Risk preference Index	Categorical variable summarizing the different risk aversion categories (highest value 6=extreme risk aversion and lowest value 1= neutral risk preference). See Table 3
Instrument (for risk index)	
Security of tenure	Whether the landlord expects an increase in the land size in the coming five years (1=increase; 0 otherwise)
Rainfall variables	
Annual rainfall	Sum of the monthly rainfall observations
Rainfall variability	Coefficient of variation of the monthly rainfall observations: ratio of the mean to the variance of monthly rainfall
Socioeconomic and Physical Farm Characteristics of the household	
Gender head	A dummy variable representing the gender of the household head (1=female;0=male)
Age head	Head's age (years)
Education head	Head's formal education (1=read and write; 2= read only; 3=none)
Male labor	The number of male working-age family members of the household
Female labor	The number of female working-age family members of the household
Oxen	The number of oxen of the household
Livestock	The number of livestock of the household
Farm size	Total farm size (ha)
Fertile plots	Average number of fertile plots
Flat plots	Average number of flat slopped plots
Distance	Approximate plot distance from homestead in meters

Table 2: Summary Statistics

Variables	2005		2007	
	Mean	Std. Dev.	Mean	Std. Dev.
Diversity	4.100	2.122	4.244	1.934
Neutral risk aversion	0.010	0.101	0.015	0.121
Slight risk aversion	0.017	0.130	0.006	0.079
Intermediate risk aversion	0.160	0.367	0.232	0.422
Moderate risk aversion	0.096	0.295	0.061	0.239
Severe risk aversion	0.245	0.430	0.218	0.413
Extreme risk aversion	0.490	0.500	0.469	0.499
Risk preference Index	4.079	1.111	3.882	1.276
Security of tenure	0.185	0.389	0.080	0.271
Annual rainfall	1041.001	219.036	1503.631	172.930
Rainfall variability	1.180	0.147	1.196	0.098
Gender head	0.173	0.378	0.191	0.393
Age head	49.853	15.567	51.229	14.956
Education head	0.386	0.487	0.333	0.471
Male labor	3.113	1.669	1.861	0.981
Female labor	2.963	1.453	1.644	0.848
Oxen	2.992	1.840	2.001	0.072
Livestock	4.259	3.222	4.137	3.138
Farm size	1.167	0.783	1.481	1.187
Fertile plots	2.155	2.130	2.771	2.475
Flat plots	0.757	0.332	0.712	0.288
Distance	70.146	51.948	72.483	53.271

Table 3: Choice sets for the risk preference experiment and risk index

	Bad harvest	Good harvest	Expected mean	Spread	Risk Aversion Category	Risk Index
Choice Set 1	100	100	100	0	Extreme risk aversion	6
Choice Set 2	90	180	105	90	Severe risk aversion	5
Choice Set 3	80	240	160	160	Moderate risk aversion	4
Choice Set 4	60	300	180	240	Intermediate risk aversion	3
Choice Set 5	20	360	190	360	Slight risk aversion	2
Choice Set 6	0	400	200	400	Neutral risk aversion	1

Table 4: Determinants of Diversity

	Naive-Poisson random effects		Pseudo-fixed effects		IV GLM (QVF)	
	Coefficient	Std. Error	Coefficient	Std. Error	Coefficient	Std. Error
Risk preference Index	0.011	(0.007)	0.01	(0.007)	0.372**	(0.173)
Annual rainfall	0.000***	(0.000)	0.000***	(0.000)	0.001***	(0.000)
Rainfall variability	0.270**	(0.083)	0.375***	(0.085)	0.734**	(0.286)
Gender head	-0.104***	(0.030)	-0.086**	(0.030)	-0.276***	(0.092)
Age head	0	(0.000)	0	(0.000)	-0.002	(0.002)
Education head	0.018	(0.020)	0.016	(0.020)	0.115	(0.074)
Male labor	0.025	(0.020)	0.002	(0.025)	0.152**	(0.073)
Female labor	0.002	(0.022)	-0.007	(0.027)	0.074	(0.076)
Oxen	0.044***	(0.007)	0.017	(0.011)	0.185***	(0.024)
Livestock	0.000*	(0.000)	0	(0.000)	0.001**	(0.000)
Farm size	0.222***	(0.024)	0.118***	(0.028)	0.829***	(0.109)
Fertile plots	0.051***	(0.004)	0.052***	(0.004)	0.232***	(0.017)
Flat slope plots	-0.033	(0.030)	-0.044	(0.030)	-0.116	(0.106)
Male labor^2	-0.003	(0.003)	-0.002	(0.003)	0.013	(0.010)
Female labor^2	0.002	(0.003)	0.003	(0.003)	0.007	(0.011)
Oxen^2	0.005	(0.024)	-0.001	(0.024)	-0.131	(0.085)
Farm size^2	-0.018***	(0.005)	-0.019***	(0.005)	-0.051*	(0.028)
Average farm size			0.149***	(0.020)		
Average oxen			0.030*	(0.014)		
Average livestock			0	(0.000)		
Average age			-0.001	(0.002)		
Average male			0.015	(0.016)		
Average female			-0.002	(0.017)		
Constant	0.051	(0.165)	-0.122	(0.174)	1.096	(0.748)
Number of observations	3069		3069		3069	
Wald Chi2	349.25		479.43		349.25	
P-Value: Prob>Chi2	0.00		0.00		0.00	
Chibar2	18.03		16.48			
P-Value: Prob>Chibar2	0.00		0.00			
Amemiya-Lee-Newey Chi2 (p-value)					0.52	
Wald test of exogeneity (p-value)					0.00	
Sargan statistic (p-value)					0.40	

The dependent variable is Diversity