WORKING PAPERS

N° 1266

November 2021

"Fertilizer use and risk: New evidence from Sub-Saharan Africa"

Céline Nauges, Douadia Bougherara and Estelle Koussoubé

Toulouse School of Economics

Fertilizer use and risk: New evidence from Sub-Saharan Africa

Céline Nauges (corresponding author)
Toulouse School of Economics, INRAE, University of Toulouse Capitole, Toulouse, France
celine.nauges@inrae.fr

Douadia Bougherara
CEE-M, Univ. Montpellier, CNRS, INRAE, Institut Agro, Montpellier, France
douadia.bougherara@inrae.fr

Estelle Koussoubé
Africa Gender Innovation Lab, World Bank, Washington, DC, USA
mkoussoube@worldbank.org

Abstract

Using a large representative dataset of 4,428 maize farmers from Burkina Faso with information on over 7,800 plots, we study the role of risk and farmers' risk preferences in their use of nitrogen fertilizers. After characterizing the role of nitrogen on the moments of the maize yield distribution, we plug the plot-specific distributions into a structural model that allows for both risk preferences and probability distortion to elicit farmers' underlying behavioural model. We found farmers to be only moderately risk averse and to distort probabilities; i.e., farmers overweight the small probabilities of getting high yields. Finally, running simulations, we find that prices are a more important driver of the quantity of nitrogen used on maize plots than farmers' risk preferences. Our results suggest that input subsidy programs in this context, if well implemented, may have the potential to increase fertilizer use.

Acknowledgments: This work was publicly funded through ANR (the French National Research Agency) under the "Cultiver et Protéger Autrement" program with the reference 20-PCPA-0005. Céline Nauges also acknowledges funding from the ANR under the Investments for the Future (Investissements d'Avenir) program, grant ANR-17-EURE-0010. Estelle Koussoubé acknowledges funding of the World Bank Group's Umbrella Facility for Gender Equality.

Fertilizer use and risk: New evidence from Sub-Saharan Africa

1. Introduction

Agricultural productivity is one important driver of economic growth (Fuglie and Wang, 2012). In a number of countries, agricultural productivity has been enhanced by the increasing adoption of modern inputs and improved technologies. However, and despite subsidization policies put in place in several countries to encourage the use of fertilizers, Sub-Saharan Africa is lagging behind.[1] We observe a relatively slow adoption of modern inputs and improved practices, and low agricultural productivity, which raises serious concern for future generations, especially considering the rapid demographic growth on the continent.

For many years economists have tried identifying reasons for what seems to be sub-optimal use of modern inputs and technologies or practices by farmers. A number of possible impediments have been listed, including: difficult access to credit, aversion to risk and downside risk (Dercon and Christiaensen, 2011; De Brauw and Eozenou, 2014; Emerick et al., 2016), procrastination and time preferences (Duflo, Kremer and Robinson, 2011; Le Cotty et al., 2018), low profitability of fertilizers (Duflo, Kremer and Robinson, 2008; Suri, 2011; Beaman et al., 2013), insecure property rights (Jacoby, Li and Rozelle, 2002), and lack of formal insurance markets (Mobarak and Rosenzweig, 2013; Cole et al., 2013; Karlan et al., 2014).

In this article we focus more specifically on the relationship between risk and farmers' use of fertilizers. We use a large nationally representative dataset of maize farmers with information on over 7,800 plots

[1] Recent work by Carter, Laajaj and Yang (2021) finds a large, sustained over time impact on input use and crop yields of a temporary input subsidy program offered to maize farmers in Mozambique. The subsidy benefited to both the targeted farmers and their social networks, suggesting that the program was highly cost effective. Although the results may not be generalizable to other contexts, they point to the potential of subsidies to fuel a green revolution in the African context.

managed by 4,428 individual farmers across 12 out of 13 administrative regions of Burkina Faso.[2] We study the role of risk and farmers' risk preferences in their application of nitrogen on maize plots. Observations of nitrogen applications and yields at the plot level, along with plot characteristics, allow a detailed investigation of risk-related properties of nitrogen.

Our work makes several contributions: first, using detailed plot-level information over a large sample of plots, we are able to characterize the impact of nitrogen use on the mean, variance, and skewness of the maize yield distribution while controlling for environmental conditions and plot characteristics. Estimating the impact of nitrogen on yield distribution is an important step towards our understanding of farmers' incentives to use fertilizers in a context of uncertainty. This step is often overlooked in studies on the impact of risk preferences on input use. We find evidence of non-linear relationships between nitrogen use and the first three moments of the maize yield distribution. The mean and variance of yield are found to be concave functions of nitrogen use, with nitrogen being a risk-increasing input for a relatively wide range of nitrogen levels. The maize yield distribution is found to be positively skewed with skewness being a convex function of nitrogen quantity.

Second, we plug the plot-specific yield distributions into a structural model that allows for both risk preferences and probability distortion to elicit farmers' underlying behavioural model. We test various forms of utility functions to determine the one that best fits the observed behaviour of farmers in terms of nitrogen application. Our searching procedure indicates that farmers are moderately risk averse (the relative risk aversion coefficient is estimated at 0.5) and that they distort probabilities (the probability weighting function is found to be inverse-S-shaped).

Finally, we use the structural model that best explains farmers' behaviour to simulate changes in economic conditions (nitrogen and maize prices) and changes in risk aversion on the optimal quantity of nitrogen used. Our simulation results suggest that prices are a more important driver of the quantity

[2] Burkina Faso is divided into 13 administrative regions and 45 provinces. We exclude the Sahel region (the driest region in the country) from the forthcoming analysis for the main reason that nitrogen has been applied only on one plot in this region.

of nitrogen used on maize plots than farmers' risk preferences. Therefore, fertilizer subsidies, if well implemented and well targeted, may have the potential to increase fertilizer use.

For the most part, our empirical study focuses on the Sudan-Sahelian zone, which is one of three main agro-ecological zones in Burkina Faso. This zone was chosen after checking for possible selection effects regarding the choice of plots on which nitrogen is applied. The Sudan-Sahelian zone appeared to be free from selection effects. It covers 24 out of the 45 provinces with significant heterogeneity in plot and farmer characteristics which we control for.

The rest of this paper is organized as follows. In Section 2 we review studies of the risk-related properties of inputs and the literature that covers the modelling of farmers' behaviour under risk. In Section 3, we describe the context for smallholder family farms operating in Burkina Faso and we provide some details on the data and main variables of interest. The structural model describing farmers' optimal production choices is presented in Section 4. In Section 5 we discuss the methodology and present the results. Section 6 concludes.

2. Related literature

Our work relates to literatures aiming at assessing the risks associated with the use of agricultural inputs and especially fertilizers, the measure of farmer risk preferences, and the impact of risk preferences on farmer input use.

Risk-related properties of fertilizers

It is now well admitted that chemical inputs such as fertilizers and pesticides impact not only the mean but also higher moments of crop yield distributions (variance and skewness in particular). Different approaches have been used in the literature to study the risk-related properties of chemical inputs.

Just and Pope (1979) proposed to condition the mean and variance of output on inputs by specifying a production function made of two additive components: $y = f(x) + h(x,v)$, combining a deterministic mean function $f(x)$ with a risk function $h(x,v)$ that depends on the input vector x and a random shock v.

Noticing that the above approach to production function estimation (i.e. parameterising a deterministic relationship between inputs and output and appending an error term to it) imposes restrictions on the relationship between inputs and the moments of the output distribution,[3] Antle (1983) proposed a more flexible representation of the output distribution by specifying the *moments of the output distribution* as explicit functions of inputs x and other controls z : $m_i = m_i(x,z) + \varepsilon_i$ with *i* representing moment order (mean, variance, skewness etc.). This non-parametric approach improves flexibility in terms of the representation of the output distribution but creates an incidental parameter problem since each moment function depends on a distinct set of parameters.[4] Antle's moment-based framework has been applied in a number of studies to assess the role of various covariates (including practices or technologies) on the moments of the output distribution (e.g., Di Falco and Chavas, 2009; Di Falco and Veronesi, 2014, and Huang, Wang and Wang, 2015).

Nelson and Preckel (1989) proposed to use a *conditional beta distribution* to study the role of inputs on the distribution of output, arguing that "if information for specifying a parametric probability distribution is available, then estimation of the parametric function is likely more efficient than estimation of a non-parametric model" (referring here to Antle's non-parametric moment-based approach). The beta distribution appears to be well suited to model crop yield distributions: a beta

[3] In the Just-Pope model described above, the elasticity of the third moment with respect to an input is directly proportional to the elasticity of the second moment with respect to that input (see equation 4 in Antle, 1983).
[4] Antle (2010) went a step further in terms of the flexibility in the representation of output distribution by considering that inputs may have different effects on the negative and positive tails of the distribution. He estimated *partial* moment functions which are based on deviations above and below a reference point and tested whether inputs play the same role on positive and negative partial moments. Tack, Harri and Coble (2012) proposed to combine Antle's moment-based approach and maximum entropy techniques to estimate conditional crop yield distributions.

random variable varies from 0 to a maximum upper bound and the distribution of a beta random variable can be skewed on either side. The beta distribution (and its moments) is characterised by two shape parameters which can be expressed as functions of inputs. This approach was used by Nelson and Preckel (1989) to study the response of maize yield distributions to fertilizer application using data from farms in Iowa and later by Babcock and Hennessy (1996).[5,6] In these articles, fertilizers were found to impact the mean, variance, and skewness of maize yield but directions of the impact varied depending on the sites.

Empirical evidence regarding the risk-related properties of chemical inputs is mixed and is often site-specific since crop growth conditions usually matter (Horowitz and Lichtenberg, 1994) but also different types of measurement indicators have been used (for the case of pesticides application, see e.g., Möhring et al., 2020). As far as we know, evidence on risk-related properties of inputs is coming primarily from agricultural fields from the US and Europe and is still rare for Sub-Saharan Africa. We contribute to this literature by estimating conditional beta distributions of maize yield using detailed plot-level data from Burkina Faso, along the lines of Nelson and Preckel (1989) and Babcock and Hennessy (1996).

Eliciting farmers' risk preferences

The literature on farmers' risk aversion is vast and studies assessing farmers' risk preferences are found worldwide. Methodologies to elicit farmers' risk preferences have evolved over time, with authors now showing a preference for experimental approaches. The elicitation of risk preferences from actual data on farmers' production choices has been put into question by Lence (2009) and Just and Just (2011). Simultaneous estimation of production technology and risk preferences was quite popular in the 1990s

[5] There is a large literature on the distribution of crop yields in general (outside the question of its relationship with inputs), with a pioneering and often-cited work by Day (1965), including debates around the validity of the normality assumption and the consequence of using aggregated data (e.g., Just and Weninger, 1999; Ramirez, Misra and Field, 2003).
[6] Another possible approach to estimate the distribution of crop yield conditional on input is conditional quantile regression. See Du, Hennessy and Yu (2012) and Chavas and Shi (2015), among others.

(e.g., Love and Buccola, 1991; Saha, Shumway and Talpaz, 1994; Chavas and Holt, 1996) but a number of authors including Lence, and Just and Just, have expressed reservations regarding the separate identification of farmers' risk preferences and input-output relationships with the farm data usually available to researchers.

Over the last decade and inspired by the pioneering work of Binswanger (1980), researchers rather used incentive-compatible experimental techniques commonly based on risk games (often involving choices between lotteries) to elicit farmers' risk preferences. Risk preferences are commonly measured by fitting the data to theoretical utility models proposed in the literature, such as expected utility (EU), rank-dependent utility (RDU), and cumulative prospect theory (CPT) models.[7] In some cases the elicited risk preferences are used in a second stage of the analysis to study their relationship with farmers' observed (real) production choices (adoption of innovative practices or technologies, input use etc.). We focus here on studies from Africa, in particular: studies by Humphrey and Verschoor (2004), Wik et al. (2004), Yesuk and Bluffstone (2009), Liebenehm and Waibel (2014) and De Brauw and Eozenou (2014) that use incentivized lottery choices, and Le Cotty et al. (2018) using hypothetical lottery choices. All in all African farmers are generally found to be risk averse, with risk aversion varying with wealth. When several theories are tested and compared, the findings almost always lead to the rejection of the EU model in favour of theories allowing for probability distortion (such as RDU or CPT).

Using experimental data gathered from surveys of east Ugandan farmers, Humphrey and Verschoor (2004) showed that the EU theory fails at describing farmers' decisions. Their results call for the use of generalised EU theory which involves a non-linear transformation of probabilities (such as RDU or CPT).

Yesuf and Bluffstone (2009) applied Binswanger (1980)'s experimental design to study risk aversion of households living in the Ethiopian highlands (but here choice sets were framed to reflect real farming decisions). Across the different games played by the households, one-third to two-third of households were found severely or extremely risk-averse and the level of risk aversion was found to depend on

[7] A difficult issue with cumulative prospect theory is the definition of the reference point (Barberis, 2013).

wealth and on whether farmers played gain-only games or games that involved both gains and losses. Wik et al. (2004) also applied Binswanger (1980)'s experimental design on a sample of Zambian households. They found that individuals are risk-averse and tend to get more risk averse when the size of the gamble increases.

Using field experiments with small-scale farmers from Mali and Burkina Faso, Liebenehm and Waibel (2014) found that the average farmer in their sample tends to overweight small probabilities and to underweight large probabilities (i.e. she is characterized by an inverted S-shaped probability weighting function). The average farmer was also found to be more risk averse in the gain domain and to have lower levels of loss aversion than its Asian counterparts (as reported in Tanaka, Camerer and Nguyen, 2010, and Nguyen and Leung, 2010, for Vietnam; and Liu, 2013, for Chinese farmers). The authors also found a negative correlation between wealth and risk aversion, with wealthier farmers being associated with lower levels of risk aversion.

De Brauw and Eozenou (2014) implemented a lab-in-the-field experiment involving farmers from northern Mozambique to test and compare several models of risk preferences. The experiment was contextualized (referring to the adoption of sweet potato varieties) and was presented in the form of lottery choices over hypothetical gains. They found that RDU models explain farmers' choices better than EU models, and that relative risk aversion is not constant.

Using experiments involving hypothetical payoffs, Le Cotty et al. (2018) measured risk attitudes of maize growers from Burkina Faso. They collected data from a representative sample of households in the Tuy and Mouhoun provinces in the Sudan-Sahelian agro-ecological zone. Under the assumption that their underlying model of risk preferences was a Constant Relative Risk Aversion (CRRA) utility function, they found that most farmers were risk averse, with a relative risk aversion coefficient around 0.3-0.4.

Impact of risk preferences on input use

Our work contributes also to the literature assessing the impact of risk preferences on farmers' production decisions. We restrict our attention to the articles using data from developing countries. Le Cotty et al. (2018), which is directly relevant to our purposes, found no statistically significant relationship between risk aversion and fertilizer use. Combining experimental field study and survey data, Liu and Huang (2013) showed that Chinese farmers who are more risk averse use greater quantities of pesticides. The findings regarding the relationship between risk preferences and input use are mixed and, when such a relationship is found, the magnitude of the impact of risk aversion on input use is rarely documented.[8]

3. Context and data

We first describe the context regarding fertilizer use for maize production in Burkina Faso (Section 3.1.). A description of the data follows in Section 3.2.

3.1. Fertilizer use in Burkina Faso

With estimated 80% of the population employed in the agricultural sector, agriculture is a crucial sector for the Burkinabé economy (World Bank, 2018). Despite policies to promote fertilizer adoption, chemical fertilizer use in Burkina Faso remains low. In 2018, average fertilizer use was around 18 kilograms per hectare of arable land in Burkina Faso, below the average in Sub-Saharan Africa (20 kg/ha), but increasing from an average of 9.5 kg/ha in 2008, year when the government launched

[8] Using simulated data, Bontemps, Bougherara and Nauges (2021) showed that, in most situations, risk aversion is expected to have a moderate impact on the optimal quantity of a risk-increasing or risk-decreasing input a farmer should use.

a large input subsidy program.[9] In comparison, average fertilizer use in Latin America & Caribbean and East Asia & Pacific was 171 kg/ha and 294 kg/ha respectively in 2018.[10]

With its production expanding rapidly for decades – an expansion mostly driven by an increase in cultivated areas –, maize is currently the second main staple crop produced in Burkina Faso (after sorghum). In 2019, maize production accounted for around 35% of total cereal production, up from around 19% of cereal production in the year 2000.[11] However, maize yields have been stagnating reflecting in part the low use of improved agricultural inputs. This is regardless of the input subsidy program launched by the government in 2008, and that is still ongoing. The subsidy program, which targets both food crops, mainly rice and maize, and cash crops such as cotton (Burkina's main cash crop), aims to increase modern input use, i.e., chemical fertilizers and improved seed use, and crop yields through a decrease in inputs prices (Wanzala-Mlobela, Fuentes and Mkumbwa, 2013). A recent study of the characteristics of the fertilizer supply in Burkina Faso found however that although the subsidy program benefited a considerable number of farmers,[12] fertilizer market prices in Burkina Faso remain high in comparison to prices in other neighbouring countries, and in comparison to resources available to Burkinabé farmers (Maître d'Hôtel and Porgo, 2018). Also, even at subsidized fertilizer prices, farmers in Burkina Faso – maize farmers for instance – have been found to use sub-optimal levels of fertilizer given its estimated profitability (Koussoubé and Nauges, 2017). Using the same dataset as in this paper, Koussoubé and Nauges (2017) show that at observed (2008/2009) level of maize and fertilizer prices, the lack of profitability (measured by a marginal value cost ratio) does not explain the low use of fertilizers by maize farmers. The authors hypothesize that other factors,

[9] For more details on the input subsidy programme, see MAFAP (2013).
[10] Word Development Indicators https://data.worldbank.org/indicator/AG.CON.FERT.ZS; accessed April 21, 2021.
[11] FAOSTAT http://www.fao.org/faostat/en/#data/QC/; accessed April 21, 2021.
[12] Inefficiencies in terms of the targeting of farmers and the distribution of the subsidized inputs remain however some of main impediments of the program (Maître d'Hôtel and Porgo, 2018; Siri, 2013; Wanzala-Mlobela, Fuentes and Mkumbwa, 2013).

including the shortage of chemical fertilizers (at crucial application time) and risks, may play a role in explaining maize farmers low use of fertilizers.

Now the question of what explains the low use of fertilizers by Burkinabé farmers remains crucial for policy makers in the country[13] and in other similar contexts. Although addressing supply constraints to fertilizer use is “a staple” of policy recommendations based on the available evidence (e.g., Maître d’Hôtel and Porgo, 2018), the lack of evidence on other important and highly contextual factors continue to impede policy making. In this paper, we explore one often overlooked explanation for the low use of fertilizers in African countries, i.e. risks associated with the use of fertilizers and the effects of risk preferences on fertilizer use.

To the best of our knowledge, only one previous study has explored the relationship between risk preferences and fertilizer use in the context of Burkina Faso. Le Cotty et al. (2018) rely on field experiments to measure farmers risk and time preferences in two cotton producing regions. Combining these experimental data with survey data, the study does not find any statistically significant association between risk aversion and fertilizer use. The contribution of our study with respect to Le Cotty et al. (2018) is threefold: first, our sample has a wider geographical coverage; second, we make a detailed analysis of the impact of nitrogen use on the distribution of maize yield (up to the third moment) while controlling for heterogeneity and selection effects, and third, we assess the role of risk preferences on nitrogen use using a structural model of farmers’ behaviour.

3.2. Data and descriptive statistics

In this study, we use a large nationally representative dataset of maize farmers in Burkina Faso, collected during the second phase of Burkina Faso’s Agricultural Census (Recensement Général Agricole 2008/2009). The survey, described in Koussoubé and Nauges (2017), provides detailed

[13] See Burkina Faso’s National Economic and Social Development Plan (Plan National de Développement Economique et Social – PNDES 2016-2020).

information on agricultural inputs, soil conditions, characteristics of individual farmers and their households, access to infrastructure and markets, and other key variables.

In the data, average maize yield was around 1,300 kg/ha, with nitrogen applied to around one-third of the plots. For those plots, the average nitrogen application was 34 kg/ha and the median was 26 kg/ha. Table 1 shows the distribution of our sample of plots across Burkina's thirteen regions, along with statistics on maize yields, with and without nitrogen application, at the region level.

Table 1. Share of plots with (w/) and without (w/o) nitrogen (N) and average yield in each region

Region	Rainfall (mm)	Total # of plots	% Plots w/N	Yield w/N (kg/ha)	% Plots w/o N	Yield w/o N (kg/ha)
BOUCLE DU MOUHOUN	779	935	58%	1,495	42%	1,122
CASCADES	1061	441	68%	1,833	32%	1,435
CENTRE	656	180	20%	1,383	80%	919
CENTRE EST	737	825	27%	1,498	73%	1,357
CENTRE NORD	583	464	11%	1,128	89%	1,065
CENTRE OUEST	740	595	47%	1,699	53%	1,143
CENTRE SUD	693	640	29%	1,340	71%	1,173
EST	663	911	17%	1,429	83%	1,426
HAUTS-BASSINS	953	854	75%	1,926	25%	1,717
NORD	607	268	30%	1,040	70%	1,015
PLATEAU CENTRAL	633	605	11%	1,326	89%	1,056
SAHEL	*482*	*78*	*1%*	*800*	*99%*	*658*
SUD OUEST	853	1,083	18%	1,197	82%	764

Notes: N stands for Nitrogen. In the Sahel region, only one plot received nitrogen.

The heterogeneity in environmental conditions partly explains the variation in average yield across the 13 regions. For plots on which no nitrogen was applied, the average yield varies from 658 kg/ha in the Sahel region to 1,717 kg/ha in the Hauts-Bassins region. For plots on which some nitrogen was applied and if we exclude the Sahel region where only one plot in our sample received nitrogen, the average yield varies from 1,040 kg/ha in the Nord to 1,926 kg/ha in the Hauts-Bassins. Practices also vary significantly since the proportion of maize plots on which some nitrogen is applied varies from 11% in the Centre Nord and Plateau Central, to 75% in the Hauts-Bassins.

To account for heterogeneity in environmental conditions, we distinguish the three main agro-ecological zones (see Appendix A for a map and the classification of all the provinces into the three zones): the *Sahelian* zone (arid land in the north), the *Sudan-Sahelian* zone (arid savannah which extends through much of the central part of the country), and the *Sudanese* zone (savannah in the south and west). The Sahelian zone receives less rainfall than the other two zones on average: over the growing season in 2008/2009, the maize parcels in the Sahelian zone received 1,586 mm, while average rainfall was 2,069 mm in the Sudanese zone and 1,828 mm in the Sudan-Sahelian zone. The average plot characteristics (sample mean over all plots within each zone) are reported in Table 2.

Table 2. Key summary statistics by agro-ecological zone

	Sahelian Zone mean (sd)	Sudan-Sahelian Zone mean (sd)	Sudanese Zone mean (sd)
Average yield on all plots (kg/ha)	1,082 (603)	1,348 (691)	1,305 (872)
Average yield on plots with N (kg/ha)	1,115 (633)	1,549 (673)	1,721 (882)
Average yield on plots with no N (kg/ha)	1,076 (597)	1,245 (677)	972 (705)
Average N on all plots (kg/ha)	3 (13)	11 (24)	14 (24)
Average N on plots with N (kg/ha)	18 (26)	34 (31)	32 (26)
Maize was grown on the plot the year before (0/1)	0.93 (0.25)	0.72 (0.45)	0.60 (0.49)
Only maize on plot (0/1)	0.91 (0.29)	0.87 (0.34)	0.63 (0.48)
Plot was borrowed or rented (0/1)	0.32 (0.47)	0.42 (0.49)	0.45 (0.50)
Plot is located in the plains (0/1)	0.91 (0.29)	0.92 (0.27)	0.87 (0.33)
Soil of his/her parcel is of low quality (0/1)	0.26 (0.44)	0.27 (0.44)	0.30 (0.46)
Manure on plot (0/1)	0.78 (0.41)	0.51 (0.50)	0.25 (0.43)
Rainfall over the growing season (mm)	1,586 (75)	1,828 (153)	2,069 (215)
Total number of plots	842	4,618	2,341
Number of plots with N (% of total plots)	142 (17%)	1,568 (34%)	1,042 (45%)
Number of plots with no N (% of total plots)	700 (83%)	3,050 (66%)	1,299 (55%)

The average yield on parcels with no (chemical) nitrogen varies between 972 and 1,245 kg/ha across the three zones, with the highest average being observed in the Sudan-Sahelian zone. The latter is probably explained by the combination of relatively abundant rainfall (compared to the Sahelian zone) and the use of manure on about half of the plots. When nitrogen is applied, the highest average yield is recorded in the Sudanese zone (1,721 kg/ha) followed by the Sudan-Sahelian zone (1,549 kg/ha). The Sahelian zone has the least favourable agro-climatic conditions for maize growing. Even when nitrogen is applied, the average yield remains much lower (1,115 kg/ha on average) than in the other two agro-ecological zones.[14] In the Sahelian zone, nitrogen is used on 17% of the plots only, but manure is spread over close to 80% of the plots (this is the zone where animal farming is the most common). Crop rotation and intercropping are also quite rare in this zone. Crop rotation and intercropping are more often practiced in the other two zones. In the Sudanese zone, nitrogen and manure are applied on 45% and 25% of the plots, respectively. In the Sudan-Sahelian zone, one-third of the plots received nitrogen and 51% received manure. Figures 1 and 2 show the average quantity of nitrogen (kg/ha) applied on plots (for those plots which received nitrogen) and the average maize yield (kg/ha) for each of the 45 provinces, as computed from the data.[15]

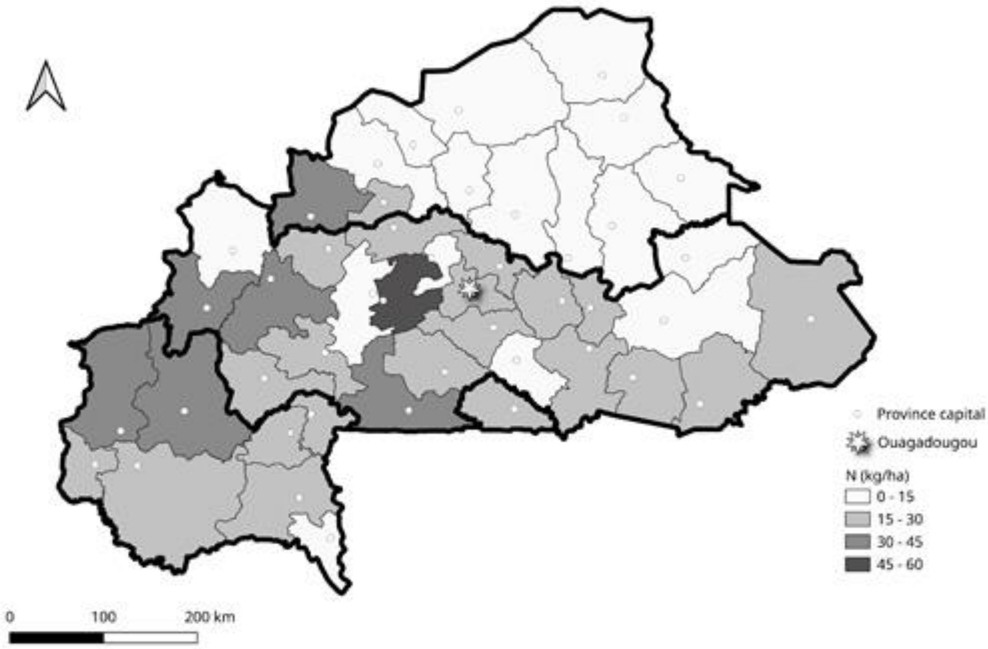


Figure 1: Weighted average of nitrogen quantity used on plots (kg/ha)

[14] The average quantity of nitrogen used on plots in the Sahelian region remains also much lower than the quantity of nitrogen used on plots in the other agro-ecological zones. See Figure 1.

[15] When computing the average quantity of nitrogen and average maize yield, weights corresponding to each plot size have been used.

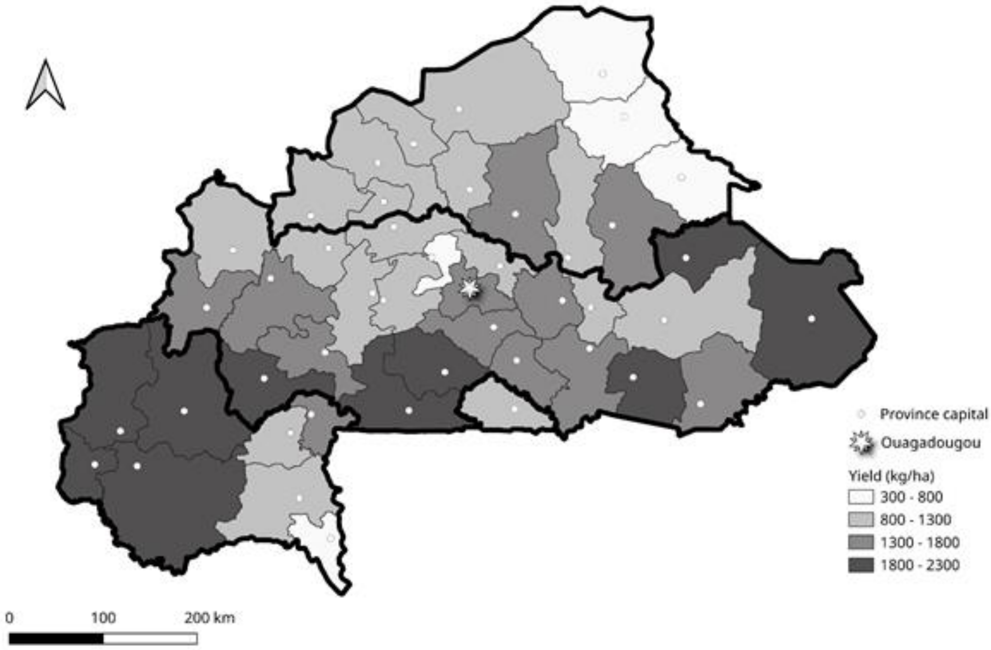


Figure 2: Weighted average of maize yield (kg/ha)

4. Farmers' optimal choice of nitrogen: conceptual model

After defining the plot-specific profit function (section 4.1.), we present in section 4.2 the structural model of farmers' preferences and behaviour that describes their optimal choice of nitrogen quantity on each plot. We assume the uncertainty faced by farmers is attached to maize yield only. Since our main variable of interest is the quantity of nitrogen used on maize plots, it is important to first characterize the relationship between nitrogen quantity and the moments of the maize yield distribution. If farmers are not risk-neutral and if nitrogen impacts the variance and skewness of the maize yield distribution, then farmers will choose nitrogen not only to increase mean yield but also to manage production risk (i.e., to reduce variance if farmers are risk averse).

4.1. Assumptions and definition of plot-specific profit

We assume that farmers make a plot-by-plot decision in terms of N application, i.e. that the decision on how much N to apply is made separately for each plot and does not depend on how much N is used on the other plots. We do as if households are unconstrained in terms of how much N they can get, which may be a strong assumption for some farmers.[16] We also assume that the price of nitrogen and the maize price are known to the farmers and non-random,[17] so that households face production risk only. If we denote by s_i the size of plot *i*, p^y the price of output and *C(N,L)* the cost of production which depends on nitrogen (*N*) and labour (*L*), then the (random) profit on plot *i* is written as follows:

$\tilde{\pi}_i = s_i \left[p^y \tilde{y}_i \left(N_i, z_i \right) - C(N_i, L_i) \right]$, with z_i, the vector of plot characteristics. The ~ sign on *y* and π indicates that these two variables are uncertain due to production risk. Under the assumption of non-random input and output prices, each moment of profit is proportional to the corresponding moment of output.

We assume that the labour costs are primarily made of the time the household spends ploughing, sowing and harvesting, and that the time spent applying nitrogen does not represent a major share of total labour. Hence we assume total labour spent on the plot (*L*) does not depend on *N*. Under this assumption, the cost of production is assumed to depend only on how much *N* is applied on the plot. We believe this assumption is reasonable in the context of our study.[18] Hence the (restricted) profit is written as:

$\tilde{\pi}_i = s_i \left[p_i^y \tilde{y}_i \left(N_i, z_i \right) - p_i^N N_i \right]$, where p^N is the unit price of *N*.

4.2. Farmers' model of risk preferences

We assume farmers maximize the utility of profit, as defined in the previous section. Farmers' risk preferences are modelled as a CRRA utility function. CRRA utility functions imply Decreasing Absolute Risk Aversion (DARA) and aversion to downside risk, some characteristics which are commonly observed in farmers' populations (Chavas, 2004). The CRRA utility function is specified as follows:

[16] For close to 50% of the parcels in our sample, farmers report difficulties in accessing agricultural inputs.
[17] A possible extension would be to assume that nitrogen and/or maize prices are random and drawn from some specific distribution. This is beyond the scope of this paper.
[18] The database records labour (in men per day) used for ploughing/tilling, sowing, maintenance, harvesting, and transportation. Nitrogen application is included in the maintenance category which, as a whole, represents around 40% of total working time (average on the entire sample).

$U(w) = \frac{1}{1-r_r} w^{1-r_r}$ when r_r is different from 1, with w representing wealth ($w > 0$) and r_r the coefficient of relative risk aversion; $U(w) = \ln(w)$ when r_r = 1 the *logarithmic function*, with $w > 0$. Wealth in the following is defined as the sum of the (household) initial wealth w_0 and the profit on the plot.[19] The above utility function can accommodate risk aversion (corresponding to the coefficient of relative risk aversion r_r being strictly positive), risk-loving behaviour ($r_r < 0$) and risk-neutrality (r_r = 0).[20]

In recent literature using data from Africa, utility models that allow for probability distortion have been shown to explain farmers' choices better than Expected Utility (EU) models (e.g., field experiments with small-scale farmers from Mali and Burkina Faso in Liebenehm and Waibel, 2014; and lab-in-the-field experiment involving farmers from northern Mozambique in De Brauw and Eozenou, 2014). In order to allow for probability distortion, the utility function over all possible outcomes (or states) *j* is defined as follows: $\sum_j \omega\left(p_j\right) U\left(\tilde{\pi}_{ij}\right)$, with p_j the objective probability of realization of profit *j*, and $\omega(p_j)$ the decision weights that are generated using the following probability weighting function: $p^\mu/[p^\mu + (1-p)^\mu]^{1/\mu}$. When $\mu = 1$, subjective and objective probabilities are equal and the utility model reduces to the EU model. When $\mu \neq 1$, the subjective and objective probabilities differ. This is the so-called Rank-Dependent Utility or RDU model. In summary, farmers are assumed to choose the level of nitrogen *N* for plot *i* that maximizes their utility $\sum_j \omega\left(p_j\right) U\left(\tilde{\pi}_{ij}\right)$ over all possible states or realizations of profit *j*, with $\tilde{\pi}_{ij} = s_i\left[p_i^y \tilde{y}_{ij}\left(N_i, z_i\right) - p_i^N N_i\right]$.

[19] We consider as initial (certain) wealth the estimated value of the herd based on average prices in each province and ownership of six types of animals: beef, draft oxen, donkey, sheep, goat and pig, and poultry (source: Ministère de l'Agriculture et de l'Hydraulique, 2011). This source of revenues can be assumed non-random in the sense that it is insensitive to production risk that may affect maize growing.

[20] In Anderson and Dillon (1992)'s terminology: r_r = 1 for somewhat risk averse; r_r = 2 for rather risk averse; r_r = 3 for very risk averse; and r_r = 4 for extremely risk averse.

5. Methodology and estimation results

We proceed in two stages. First, we characterize the uncertainty in maize yield as represented by the random variable $\tilde{y}_i$ and we assess how nitrogen impacts the first three moments of the yield distribution. For that purpose, we estimate the conditional yield distribution (section 5.1). Second, we elicit farmers' underlying model of risk preferences by testing various competing utility models allowing for both risk aversion and probability distortion (section 5.2).

5.1. How does nitrogen impact the maize yield distribution

We assume maize yield follows a beta distribution with shape parameters α and β ($\alpha > 0$ and $\beta > 0$).

5.1.1. Estimation of the conditional yield distribution

We allow these shape parameters to be parametric functions of a set of covariates including the quantity of nitrogen (*N*) used per hectare (see Babcock and Hennessy, 1996, for a similar approach). In all that follows the shape parameters are estimated by Maximum Likelihood (ML) and we control for the fact that observations from plots belonging to the same household may not be independent. For a (standardized) random variable y_s belonging to the [0,1] interval, the Probability Density Function (PDF) of the beta distribution is as follows:

$\frac{y_s^{\alpha-1}(1-y_s)^{\beta-1}}{\mathrm{B}(\alpha,\beta)}$ where $\mathrm{B}(\alpha,\beta)=\frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}$ and $\Gamma(.)$ is the gamma function.

Note that in the following the shape parameter estimates are always for the *standard* beta distribution, that is the distribution on the [0,1] interval. Maize yield varies in our sample from 112 to 3,960 kg/ha so the variable *y* measuring yield has been normalized as follows: $y_s = \frac{y-112}{3960-112}$.

We condition α and β on covariates, with primary interest for the quantity of nitrogen used on each plot. The estimation of various specifications showed the importance of controlling for heterogeneity across regions, most likely due to varying environmental conditions (soil type and quality, plot exposure, weather conditions), so we condition the two shape parameters with dummy variables for the three agro-ecological zones. Other binary covariates include whether maize was grown on the plot the year before; whether there is only maize on plot; whether the plot was borrowed or rented; whether the plot is located in the plains; whether the household head believes that the soil of his/her parcel is of low quality, and whether some manure was used on plot.

The model was estimated on the full sample of plots, which amounts to a total of 7,801 parcels belonging to 4,428 households. Results are shown in Table 3.

Table 3. Estimated parameters of conditional beta distribution of yields

	Coeff	Std. Err.	P value
Parameter alpha			
Sahelian zone (0/1)	-0.191	0.104	0.065
Sudan-sahelian zone	Ref.	Ref.	Ref.
Sudanese zone (0/1)	-0.730	0.058	0.000
Quantity of N/ha	0.005	0.002	0.016
Squared quantity of N/ha	0.000	0.000	0.144
Maize grown on plot the year before (0/1)	0.174	0.041	0.000
Only maize on plot (0/1)	0.002	0.060	0.969
Plot borrowed or rented (0/1)	0.445	0.048	0.000
Plot located in the plains (0/1)	-0.306	0.082	0.000
Soil perceived as low quality (0/1)	0.118	0.054	0.029
Some manure used on plot (0/1)	0.086	0.034	0.013
Constant	1.745	0.103	0.000
Parameter beta			
Sahelian zone (0/1)	0.757	0.310	0.015
Sudan-sahelian zone	Ref.	Ref.	Ref.
Sudanese zone (0/1)	-1.633	0.122	0.000
Quantity of N/ha	-0.049	0.004	0.000
Squared quantity of N/ha	0.000	0.000	0.000

Maize grown on plot the year before (0/1)	0.598	0.083	0.000
Only maize on plot (0/1)	-1.301	0.222	0.000
Plot borrowed or rented (0/1)	0.880	0.089	0.000
Plot located in the plains (0/1)	-1.027	0.202	0.000
Soil perceived as low quality (0/1)	0.694	0.139	0.000
Some manure used on plot (0/1)	-0.089	0.020	0.000
Constant	5.561	0.300	0.000
Log L			2,770.49
#Observations			7,801
#Households			4,428

We find evidence of a quadratic relationship between nitrogen per hectare and the shape parameters α and β. Most of the covariates are statistically significant.

5.1.2. Testing for heterogeneity and selection effects

To separate the effects of heterogeneity in conditions from the effect of nitrogen use, we first calculate the estimates of the shape parameters within each agro-ecological zone by setting the nitrogen quantity at 0 and replacing the other covariates by their zone-specific mean. The estimated shape parameters are reported in Table 4, along with the first three empirical moments of the maize yield distribution.[21]

[21] The mean, variance and skewness are all functions of the shape parameters: the mean of the beta distribution is equal to $\alpha/(\alpha+\beta)$; the variance is $\alpha\beta/\left[(\alpha+\beta)^2(\alpha+\beta+1)\right]$, and the skewness is $\left[2(\beta-\alpha)\sqrt{\alpha+\beta+1}\right]/\left[(\alpha+\beta+2)\sqrt{\alpha\beta}\right]$.

Table 4. Conditional beta distribution – Estimated moments and estimated shape parameters at zone means and for N = 0

	Estimated shape parameters and 95% confidence intervals		
	Sahelian zone	Sudan-Sahelian zone	Sudanese zone
α and 95% CI	1.68 [1.51;1.85]	1.86 [1.74;1.97]	1.11 [1.03;1.19]
β and 95% CI	5.15 [4.61;5.69]	4.12 [3.81;4.44]	3.15 [2.88;3.41]
Mean yield (kg/ha)	1,058	1,305	1,114
Variance of yield	350,644	453,491	541,997
Skewness of yield	0.748	0.544	0.800
# of plots	842	4,618	2,341

Note: CI is for Confidence Interval.

The estimated beta distribution is positively skewed in each zone since $\alpha \neq \beta$ and $\alpha < \beta$. The estimated shape parameter α shows significant heterogeneity, varying from 1.11 in the Sudanese zone to 1.86 in the Sudan-Sahelian zone and this parameter is statistically different between the three zones. The estimated shape parameter β varies from 3.15 in the Sudanese zone to 5.15 in the Sahelian zone, and is statistically different between the three zones. The estimated mean yield (with *N* set at 0) varies from 1,058 kg/ha in the Sahelian zone to 1,305 kg/ha in the Sudan-Sahelian zone. In the Sudanese zone, the yield distribution is the most right-skewed and has the largest variance. From the above estimates of the shape parameters we can draw the estimated distributions of yields for the three zones (see Figure 3).[22]

[22] In each case, the density is computed at 100 equally-spaced points over the [0,1] scale.

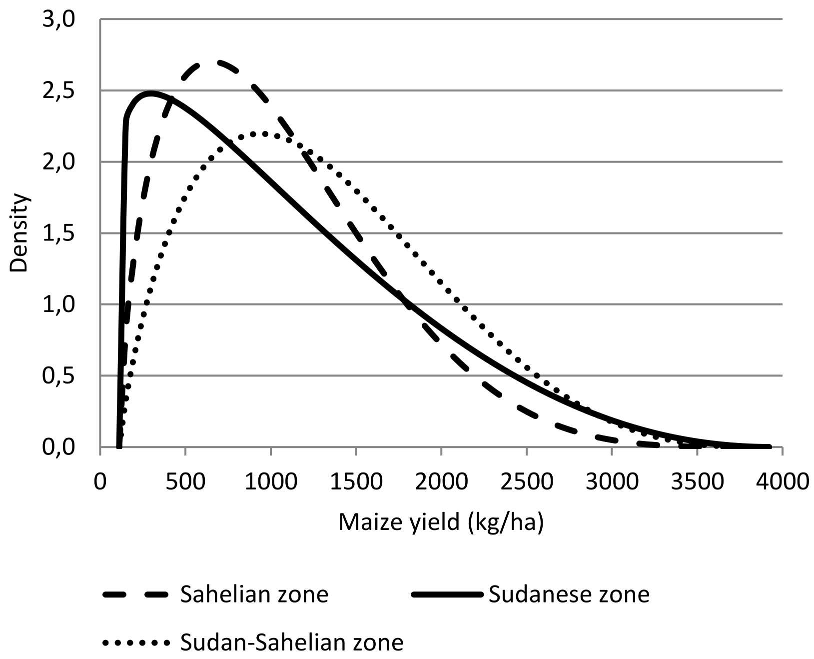


Figure 3. Estimated conditional beta distributions of maize yield for the three zones at zone means and for N = 0

Since a number of plots did not receive any nitrogen, our estimates of the shape parameters may suffer from selection bias. Selection bias would occur if the decision to apply nitrogen on a specific plot depends on unobservable characteristics which may also impact the shape parameters of the yield distribution. Several plot characteristics are used as controls in our model but there may remain unobservable characteristics (for example plot exposure) which explain both the decision to use nitrogen and the yield distribution.

To investigate a possible selection bias, we propose to compare the shape parameters of the beta distribution estimated on two sub-samples: the sub-sample of plots which did not receive any nitrogen and the sub-sample of plots on which some nitrogen was applied. If the estimated shape parameters are statistically different from one sub-sample to the other, this will indicate possible selection bias.

The analysis is shown in Appendix B.[23] The estimates raise concern in terms of a possible selection bias in the Sudanese zone since the estimated average yield, before any application of nitrogen, is about 50% higher for the plots which received nitrogen (1,454 kg/ha) compared to the average yield estimated on the plot which did not receive any nitrogen (989 kg/ha). When comparing the 95% confidence intervals of the estimated shape parameters, the Sudan-Sahelian zone appears to be less prone to selection bias than the other two zones. It is also the zone which covers the largest number of plots (4,618 in total), so we focus the rest of our analysis on the Sudan-Sahelian zone.

5.1.3. Nitrogen impact on the moments of the yield distribution

Using the estimated shape parameters from Table 4, and setting all the covariates except N at their sample mean over the Sudan-Sahelian zone, Figure 4 shows the estimated distribution of maize yield for various levels of nitrogen per hectare: no nitrogen, 10 kg/ha, 20 kg/ha, and 40 kg/ha.

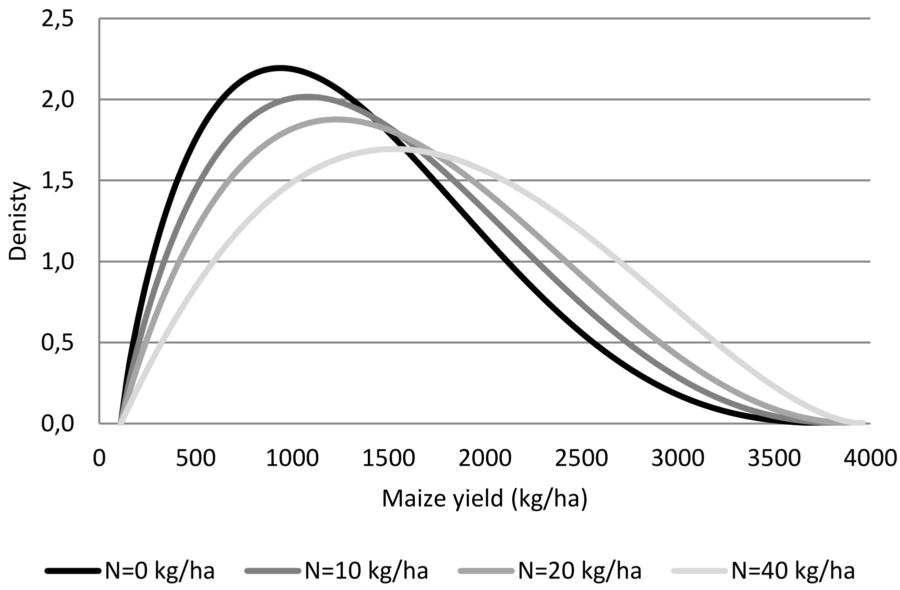


Figure 4. Estimated maize yield distribution in the Sudan-Sahelian zone for various levels of nitrogen

[23] We are not aware of any earlier studies in agriculture in which selection bias has been controlled for when estimating distributions of crop yield. There exist papers in labour economics which estimate wage distributions using quantile regression and correct for selection bias (e.g., Machado and Mata, 2005; Albrecht, van Vuuren and Vroman, 2009).

As expected, the yield distribution becomes flatter and moves to the right when the level of nitrogen applied on the plot increases. With levels above 60 kg/ha, the impact of nitrogen on the distribution of yield is found to be more limited (indeed, the three distributions almost overlap and are less skewed to the right).[24] See in Appendix C the distribution of yields with levels of nitrogen per hectare set at 60, 80 and 100 kg. This finding seems to be in line with observed practices in the sense that we observe a small number of plots (fewer than 10%) on which the average use of nitrogen is above 40 kg/ha. Finally we check how nitrogen impacts on the three moments of the maize yield distribution. Figures 5, 6, and 7 illustrate impact on the mean, variance, and skewness, respectively.

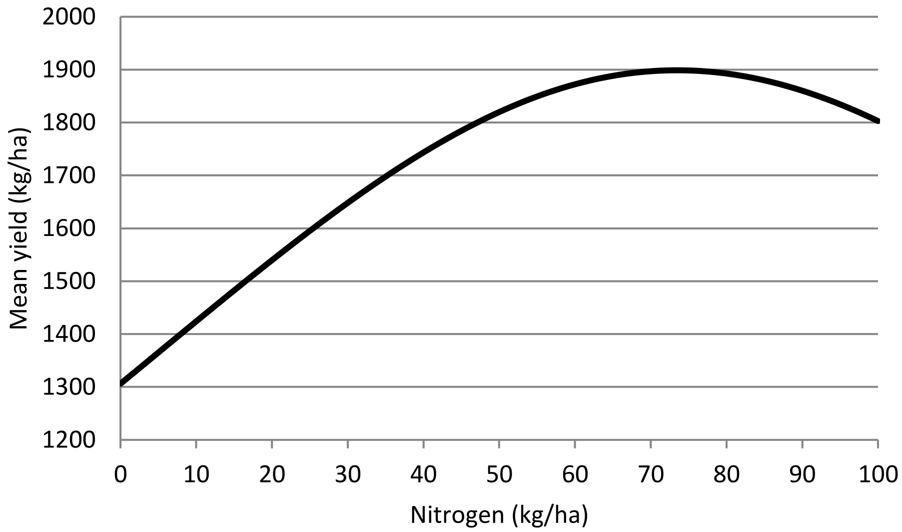


Figure 5. Impact of nitrogen use on the mean of the maize yield distribution in the Sudan-Sahelian zone

[24] Interestingly, Babcock and Hennessy (1996) get something very similar when using experimental data from Iowa farms.

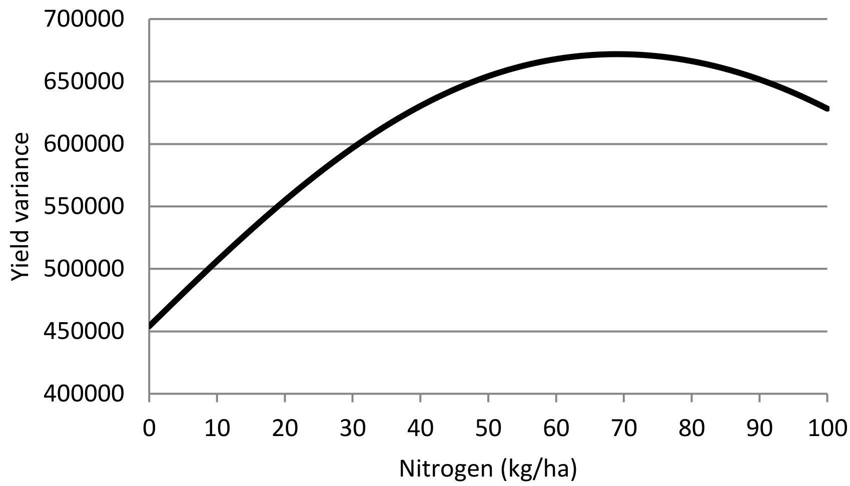


Figure 6. Impact of nitrogen use on the variance of the maize yield distribution in the Sudan-Sahelian zone

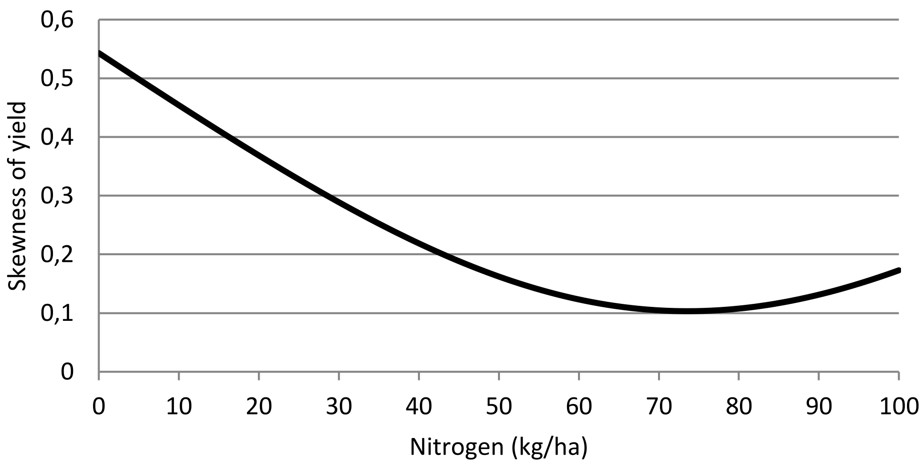


Figure 7. Impact of nitrogen use on the skewness of the maize yield distribution in the Sudan-Sahelian zone

These graphs illustrate the complex relationship between nitrogen use and the moments of the maize yield distribution. Both the mean and variance functions are concave functions of N, with the mean yield reaching its maximum when N = 73 kg/ha and the variance of yield being at its highest when N = 69 kg/ha. Skewness is positive, it decreases until N = 74 kg/ha and increases for values above 75 kg/ha. On our sample the average nitrogen quantity applied on plots is 34 kg/ha, which is much lower than the quantity of nitrogen that maximises mean yield. There are different possible reasons to explain this finding: a first possible explanation is that farmers are risk averse. Since nitrogen is found to be risk-increasing for quantities of nitrogen below 69 kg/ha (Figure 6), then risk-averse farmers will tend to use quantities of nitrogen that are below the level that maximizes the mean yield in order to reduce the variability in yield. Farmers are usually found to be averse to downside risk, which means that they dislike the risk of facing very low yields. In our case, the skewness of yield is found to be positive so farmers may be inclined to reduce their use of nitrogen in order to maintain the possibility of getting very high yields. When faced with skewed distributions, it may also become more difficult for farmers to assess objective probabilities of rare outcomes and they may overweight the probability of high yield outcome. Another possible explanation is that farmers are constrained in how much quantity of nitrogen they can get, which is consistent with the observation that farmers face difficulties in accessing inputs for about half of the parcels in our sample.[25] In this section, we assessed the impact of fertilizers on the moments of the distribution of yields, an important but often overlooked step when trying to understand farmers' level of input use. In the next section we allow for both risk aversion and probability distortion when searching for the underlying model of farmers' behaviour.

5.2. Elicitation of farmers' risk preferences: methodology and results

5.2.1. Methodology

The estimation of risk preference parameters from observed production data raises significant identification challenges (Lence, 2009; Just and Just, 2011). For this reason, we decided not to *estimate* the unknown parameters of the utility function (the coefficient of relative risk aversion, r_r) and of the decision weighting function (the distortion parameter, μ), but instead to *search* for the pair of parameters (r_r, μ) that provides the best fit to our data. After searching over a wide range of parameter values, we narrowed our selection down to three possible levels of the coefficient of relative risk aversion r_r: 0 (risk-neutrality), 0.5 (moderate risk aversion in the range of what has been

[25] Around 47% farmers also reported the high cost of agricultural inputs as a difficulty to have access to agricultural inputs.

estimated by Le Cotty et al. (2018), using data from the same country), and 2 (slightly stronger risk-aversion) and three levels for the distortion parameter (μ = 0.8, 1 and 1.2). For each combination of the two parameters and to determine the optimal *N*, we proceed as follows: using the above ML estimates of the beta distribution parameters, we calculate the estimated maize yield distribution for each parcel in the Sudan-Sahelian zone (4,618 in total) and each level of nitrogen varying from 0 to 100 (*N*=0, 1, 2 … 100). For each possible level of *N* and each plot *i*, the maize yield distribution is fully characterized by the estimated shape parameters $\hat{\alpha}$ and $\hat{\beta}$, which both depend on *N* and z_i, the set of plot characteristics. From the plot-specific yield distributions and using information on maize and nitrogen prices, we easily derive plot-specific profit distributions, which also vary with how much nitrogen is applied on each plot. For each of the 4,618 x 101 distributions, we calculate the utility of profit and the corresponding probability (objective or subjective depending on the value of μ) at 1,000 equidistant points over the [0,1] range. When μ is set equal to 1, we assume farmers use objective probabilities. When μ is different from 1, we assume farmers distort probabilities. Summing over the 1,000 observations [utility x probability], we get an estimate of the utility of profit for each parcel and each level of nitrogen from 0 to 100. For each parcel the optimal level of *N* is the one for which the utility of profit reaches its maximum.

When available, we consider the maize price and price of nitrogen as reported by the farmers in our sample (after removing outliers, below the 5th percentile and above the 95th percentile of each price distribution).[26] If the price information is missing, we use the median price in the province. If the latter is also missing, we use the median price in the region. Summary statistics on prices for the Sudan-Sahelian zone are shown in Table 5.

[26] The maize price is not subsidized but the price of nitrogen can be, in particular if farmers buy from government agencies which supply fertilizers at subsidized prices. The latter can explain the wide range of N prices in Table 5.

Table 5. Maize and nitrogen prices in the sample (CFA francs/kg)

	Mean	1st quartile	Median	3rd quartile	Min	Max
Maize price (CFA francs/kg)	126	102	117	139	53	347
N price (CFA francs/kg)	1,535	1,256	1,429	1,727	739	3,175

5.5.2. Identifying underlying parameters of farmers' preferences

We report in Table 6 some statistics to compare observed and predicted nitrogen (N) on plots for various combinations of the relative risk aversion coefficient (r_r) and distortion parameter (μ). Column 2 shows the actual number of observed plots which did not receive any nitrogen and the actual number of plots which received some nitrogen, along with the average quantity of nitrogen applied. The following three columns report the same outcomes as predicted by the RDU model for the corresponding pair of parameters. In each case we also report the sum of absolute differences between the predicted and observed quantity of nitrogen, with smaller sums indicating better fit to the observed data. Based on this statistic, the RDU model that provides the best fit to the data features a coefficient of relative risk aversion equal to 0.5 and a distortion parameter equal to 0.8.

Table 6. Predictions from RDU model and goodness of fit

		Predictions with RDU model		
Risk-neutral (r_r = 0)	Observed plots	$r_r = 0$ $\mu = 0.8$	$r_r = 0$ $\mu = 1$	$r_r = 0$ $\mu = 1.2$
# plots without N	3,059	3,153	2,986	2,933
# plots with N	1,559	1,465	1,632	1,685
Mean N if N>0 (kg/ha)	34	33	35	35
Sum of absolute differences	-	77,748	82,945	84,710
Moderately risk averse (r_r = 0.5)		$r_r = 0.5$ $\mu = 0.8$	$r_r = 0.5$ $\mu = 1$	$r_r = 0.5$ $\mu = 1.2$
# plots without N	3,059	3,227	3,028	2,966
# plots with N	1,559	1,391	1,590	1,652
Mean N if N>0 (kg/ha)	34	32	34	35
Sum of absolute differences	-	76,816	81,878	83,835
Rather risk averse (r_r = 2)		$r_r = 2$	$r_r = 2$	$r_r = 2$

		μ = 0.8	μ = 1	μ = 1.2
# plots without N	3,059	3,244	3,010	2,968
# plots with N	1,559	1,374	1,608	1,650
Mean N if N>0 (kg/ha)	34	39	40	40
Sum of absolute differences	-	84,956	89,947	91,536

A coefficient of relative risk-aversion of 0.5 is in line with estimates obtained by Le Cotty et al. (2018) on farmers from Burkina Faso (using experiments, they estimated the risk aversion coefficient of farmers to be in the range 0.3-0.4). A distortion parameter lower than 1 indicates that the probability-weighting function is inverse S-shaped, with low probability events being over-weighted and high probability events being under-weighted. Inverse-S-shaped probability weighting functions are rather common among farmers (see Appendix 7 in Bontemps, Bougherara and Nauges, 2021 for a review). In the following we look more closely at the predictions from this particular model.

5.2.3. Assessing the role of risk aversion and economic conditions on nitrogen use

In what follows we use the best-fitted model (RDU with relative risk aversion coefficient set at 0.5 and probability distortion parameter set at 0.8) to further investigate the role of risk aversion and economic conditions on optimal use of nitrogen.

The role of prices

We study how the optimal *N* varies with the price of nitrogen and with the maize price in the best performing model, identified above. We consider an increase in the price of maize (for each plot) by 20% and 40%, compared to the actual observed price, and two other scenarios assuming a 20% and a 40% decrease in the price of nitrogen (for each plot). These variations (20 to 40%) are within the observed cross-sectional variation in prices on our sample (see Table 5). Simulation results are shown in Tables 7 and 8.

Table 7. Predicted N applications with higher maize prices; RDU Model (r_r = 0.5, μ = 0.8)

Range of N application (kg/ha)	Observed prices, benchmark		20% increase in maize price		40% increase in maize price	
	# of plots	Share of plots	# of plots	Share of plots	# of plots	Share of plots
0	3,227	70%	2,208	48%	1,441	31%
[1 to 10]	121	3%	110	2%	124	3%
[11 to 20]	154	3%	237	5%	355	8%
[21 to 30]	310	7%	675	15%	437	9%
[31 to 50]	684	15%	1,152	25%	1,765	38%
51 and above	122	3%	236	5%	496	11%
Total	4,618	100%	4,618	100%	4,618	100%
Mean N if N>0 (kg/ha)	32		34		37	

When the price of maize increases, it becomes more profitable to use fertilizers on plot so the number of plots for which it is optimal to use some nitrogen increases (Table 7). In the benchmark case (RDU model using observed prices), it is optimal not to apply any nitrogen on 70% of the plots. If the price of maize increases by 20%, then it is optimal not to use any nitrogen on 48% of the plots. If the price of maize increases by 40%, only 31% of the plots should not receive any nitrogen. The average nitrogen application on plots for which the optimal *N* is greater than 0 increases from 32 kg/ha on average in the benchmark case to 34 kg/ha with a 20% increase in maize price and to 37 kg/ha with a 40% increase in maize price.

We report in Table 8 the optimal *N* application when the price of nitrogen decreases by 20% and 40% compared to the actual prices, as predicted by the best-fitted RDU model. It is worth noting that a decrease of 20-40% corresponds roughly to the subsidy levels offered by the government of Burkina Faso to farmers during the agricultural season 2008/2009.[27] We find that when the price of nitrogen decreases by 20%, the number of plots on which it is optimal not to apply any nitrogen decreases to

[27] Koussoubé and Nauges (2017, pp. 202-203) quote a World Bank report indicating that "the price of nitrogen in Burkina Faso was estimated at 1,961 CFA franc/kg without any subsidies while the subsidised price was estimated at 1,156 CFA franc/kg". The difference between the two corresponds roughly to a 40% subsidy.

45%. With a 40% decrease in nitrogen price, the share of plots on which no nitrogen should be applied is 16%. Under this scenario, the optimal *N* lies in the range 31 to 50 kg/ha for close to 50% of the plots, compared to 15% in the benchmark scenario. As a consequence the average quantity of N that is applied increases from 32 kg/ha in the benchmark case to 41 kg/ha when the price of N decreases by 40% (a 28% increase). Although observed average nitrogen prices in our dataset partially reflect the subzidized price available to some farmers of our sample, these results suggest that the government input subsidy program implemented in 2008/2009 had the potential to substantially increase the use of nitrogen among maize farmers in the study area.

Table 8. Predicted N applications with lower nitrogen prices; RDU Model (r_r = 0.5, μ = 0.8)

	Observed prices, benchmark		20% decrease in nitrogen price		40% increase in nitrogen price	
Range of N application (kg/ha)	# of plots	Share of plots	# of plots	Share of plots	# of plots	Share of plots
0	3,227	70%	2,066	45%	752	16%
[1 to 10]	121	3%	92	2%	120	3%
[11 to 20]	154	3%	224	5%	205	4%
[21 to 30]	310	7%	507	11%	410	9%
[31 to 50]	684	15%	1,431	31%	2,170	47%
51 and above	122	3%	298	6%	961	21%
Total	4,618	100%	4,618	100%	4,618	100%
Mean N if N>0 (kg/ha)	32		36		41	

The role of risk aversion

We check by how much risk aversion changes the optimal level of nitrogen by comparing predictions in terms of N application based on the RDU best-fitted model and a RDU model assuming risk neutrality (r_r = 0) while keeping the distortion parameter the same (μ = 0.8). Results are shown in Table 9.

Table 9. Predicted N applications with and without risk aversion; RDU Model (μ = 0.8)

	RDU model, benchmark		RDU model + risk neutrality (r_r = 0)	
Range of N application (kg/ha)	# of plots	Share of plots	# of plots	Share of plots
0	3,227	70%	3,153	68%
[1 to 10]	121	3%	129	3%
[11 to 20]	154	3%	154	3%
[21 to 30]	310	7%	321	7%
[31 to 50]	684	15%	735	16%
51 and above	122	3%	126	3%
Total	4,618	100%	4,618	100%
Mean N if N>0 (kg/ha)	32		33	

The level of the relative risk aversion coefficient has a relatively small impact on the number of plots within each range of N application. Under risk neutrality the percentage of plots which do not receive any nitrogen decreases to 68%, compared to 70% in the model with risk aversion. The average N application is almost unchanged when assuming risk neutrality.

Under the assumption that our measure of farmers' risk aversion reflects the truth, then the simulation results show that the optimal quantity of nitrogen used is more sensitive to changes in prices than to the level of risk aversion. Consequently, input subsidization policies are likely to be more effective in incentivizing farmers to use more nitrogen on plots than, for example, insurance policies. Our results confirm those of Le Cotty et al. (2018) who found that risk aversion (elicited using experimental techniques) had no significant impact on farmers' use of fertilizers in two provinces of Burkina Faso. They are also in line with simulation results in Bontemps, Bougherara and Nauges (2021) showing that prices are, in most cases, a more important driver of input use than risk preferences.

6. Conclusion

Using a large and representative sample of maize plots from Burkina Faso, we provide new evidence on the role risk may play in explaining low levels of fertilizer use in Sub-Saharan Africa. We find that nitrogen use impacts the first three moments of the maize yield distribution and that nitrogen is risk-increasing for a wide range of nitrogen levels. However, the risk aversion of maize growers is found to be rather moderate on our sample and input use under risk aversion is not very different from the predicted optimal input use if farmers were risk-neutral. Risk thus does not appear to be the main reason for explaining low levels of fertilizer use on maize plots.

These findings contribute to the literature assessing the impact of risk on farmers' production decisions. We provide a quantitative assessment of the expected change in fertilizer use induced by risk aversion, and we compare this expected impact to those induced by fertilizer and maize price changes. Optimal input use is found to be sensitive to changes in nitrogen and maize prices, arguing in favour of a continuous subsidization of fertilizer prices. Simulations of input use for prices in the range of subsidised prices provided by the government suggest that the government's subsidy program could have the expected impacts (at least at short term) on input use and yield, provided an improvement in the implementation of the program. For instance, a decrease in observed nitrogen prices between 20-40% increases the profitability of fertilizer use for both risk-neutral and moderately risk-averse farmers. However, farmers reports suggest that a large portion of farmers face important constraints in accessing agricultural inputs (at market or subsidized prices). Also, besides targeting issues that impede the program efficiency, important administrative and logistical costs associated with program implementation affect the cost-effectiveness of the program. [28]

Several caveats are in order. First, it was not possible to directly observe and control for the possible supply constraints farmers may face, either in relation to fertilizer quantity or quality.[29] Our analysis was run under the assumption that prices are known to the farmers, which may be a rather strong assumption in some cases. We did our best to control for heterogeneous conditions and plot characteristics but there may remain important omitted variables. We studied farmers' use of nitrogen considering only maize plots. One reason that may explain the relative moderate role of risk aversion

[28] For an analysis of costs associated with the implementation of the subsidy and factors impeding the efficacity of the program, see for example Maître d'Hôtel and Porgo (2018) and Siri (2013).

[29] A number of farmers who reported difficulties in accessing inputs still used some nitrogen on their plots.

is that revenues from maize production may represent a small share of farms' income and wealth, especially for those households who own animals. Finally, the elicitation of risk preferences is based on the estimated maize yield distribution. The distribution of maize yield that farmers perceive or have in mind may be different from the one that has been estimated in this paper.

References

Albrecht, J., van Vuuren, A., and Vroman, S. (2009). Counterfactual distributions with sample selection adjustments: Econometric theory and an application to the Netherlands. Labour Economics 16(4): 383-396.

Anderson, J.R., and Dillon, J.L. (1992). Risk analysis in dryland farming systems. Farming systems management series no. 2, Food and Agriculture Organization of the United Nations (FAO). FAO, Rome.

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

Antle, J.M. (2010). Asymmetry, partial moments and production risk. American Journal of Agricultural Economics 92(5): 1294-1309.

Babcock, B.A., and Hennessy, D.A. (1996). Input demand under yield and revenue insurance. American Journal of Agricultural Economics 78(2): 416-427.

Barberis, N. C. (2013). Thirty years of prospect theory in economics: A review and assessment. Journal of Economic Perspectives 27(1): 173-196.

Beaman, L., Karlan, D., Thuysbaert, B., and Udry, C. (2013). Profitability of fertilizer: Experimental evidence from female rice farmers in Mali. American Economic Review 103(3): 381-386.

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

Bontemps, C., Bougherara, D., and Nauges, C. (2021). Do risk preferences really matter? The case of pesticide use in agriculture. Environmental Modeling & Assessment 26: 609-630.

Carter, M., Laajaj, R., and Yang, D. (2021). Subsidies and the African green revolution: Direct effects and social network spillovers of randomized input subsidies in Mozambique. American Economic Journal: Applied Economics 13(2): 206-229.

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

Day, R.H. (1965). Probability distributions of field crop yields. Journal of Farm Economics 47(3): 713-741.

De Brauw, A., and Eozenou, P. (2014). Measuring risk attitudes among Mozambican farmers. Journal of Development Economics 111: 61-74.

Dercon, S., and Christiaensen, L. (2011). Consumption risk, technology adoption and poverty traps: Evidence from Ethiopia. Journal of Development Economics 96(2): 159-173.

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.

Di Falco, S., and Veronesi, M. (2014). Managing environmental risk in presence of climate Change: The role of adaptation in the Nile Basin of Ethiopia. Environmental and Resource Economics 57: 553-577.

Duflo, E., Kremer, M., and Robinson, J. (2008). How high are rates of return to fertilizer? Evidence from field experiments in Kenya. American Economic Review 98(2): 482-488.

Duflo, E., Kremer, M., and Robinson, J. (2011). Nudging farmers to use fertilizer: Theory and experimental evidence from Kenya. American Economic Review 101(6): 2350-2390.

Emerick, K., de Janvry, A., Sadoulet, E., and Dar, M. H. (2016). Technological innovations, downside risk, and the modernization of agriculture. American Economic Review 106(6): 1537–1561.

Fuglie, K., and Wang, S. (2012). Productivity growth in global agriculture shifting to developing countries. Choices 27: 4.

Horowitz, J. K., and Lichtenberg, E. (1994). Risk-reducing and risk-increasing effects of pesticides. Journal of Agricultural Economics 45(1): 82-89.

Huang, J., Wang, Y., and Wang, J. (2015). Farmers' adaptation to extreme weather events through farm management and its impacts on the mean and risk of rice yield in China. American Journal of Agricultural Economics 97(2): 602-617.

Humphrey, S., and Verschoor, A. (2004). Decision making under risk among small farmers in East Uganda. Journal of African Economies 13(1): 44-101.

Jacoby, H.G., Li, G., and Rozelle, S. (2002). Hazards of expropriation: Tenure insecurity and investment in rural China. American Economic Review 92(5): 1420-1447.

.Just, R. E., and Just, D. R. (2011). Global identification of risk preferences with revealed preference data. Journal of Econometrics 162(1): 6-17.

Just, R.E., and Pope R.D. (1979). Production function estimation and related risk considerations. American Journal of Agricultural Economics 61(2): 277-284.

Just, R.E., and Weninger, Q. (1999). Are crop yields normally distributed? American Journal of Agricultural Economics 81(2): 287-304.

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

Le Cotty, T., Maître d'Hôtel, E., Soubeyran, R., and Subervie, J. (2018). Linking risk aversion, time preference and fertiliser use in Burkina Faso. The Journal of Development Studies 54(11): 1991-2006.

Lence, S. H. (2009). Joint estimation of risk preferences and technology: Flexible utility or futility? American Journal of Agricultural Economics 91(3): 581-598.

Liebenehm, S., and Waibel, H. (2014). Simultaneous estimation of risk and time preferences among small-scale cattle farmers in West Africa. American Journal of Agricultural Economics 96(5): 1420-1438.

Liu, E.M. 2013. Time to change what to sow: Risk preferences and technology adoption decisions of cotton farmers in China. Review of Economics and Statistics 95 (4): 1386-1403.

Liu, E.M., and Huang, J. (2013). Risk preferences and pesticides use by cotton farmers in China. Journal of Development Economics 103: 202-215.

Love, H.A., and Buccola, S.T. (1991). Joint risk preference-technology estimation with a primal system. American Journal of Agricultural Economics 73(3): 765-774.

Machado, J.A.F., and Mata, J. (2005). Counterfactual decomposition of changes in wage distributions using quantile regression. Journal of Applied Econometrics 20: 445-465.

MAFAP (2013). Review of food and agricultural policies in Burkina Faso. MAFAP Country Report Series. Rome: FAO.

Maître d'Hôtel, E. and Porgo, I. (2019). Quelles actions politiques pour renforcer l'accès des producteurs agricoles à des engrais chimiques à prix réduits au Burkina Faso ? Rapport d'analyse de politique, SAPAA (Projet de Suivi et analyse des politiques agricoles et alimentaires). Rome, FAO. 54 pages. *[In French]*

Ministère de l'Agriculture et de l'Hydraulique (2011). Rapport général du module élevage. Phase 2 RGA 2008. Bureau central du recensement général de l'agriculture. *[In French]*

Mobarak, A. M., and Rosenzweig, M.R. (2013). Informal risk sharing, index insurance, and risk taking in developing countries. American Economic Review 103(3): 375-380.

Möhring, N., Bozzola, M., Hirsch, S., and Finger, R. (2020). Are pesticides risk decreasing? The relevance of pesticide indicator choice in empirical analysis. Agricultural Economics 51(3): 429-444.

Nelson, C., and Preckel, P. (1989). The conditional beta distribution as a stochastic production function. American Journal of Agricultural Economics 71(2): 370-378.

Nguyen, Q., and Leung, P. (2010). How nurture can shape preferences: An experimental study on risk preferences of Vietnamese fishers. Environment and Development Economics 15 (5): 609-631.

PNUD and Ministry of Animal Resources (2011). Contribution de l'élevage à l'économie et à la lutte contre la pauvreté, les déterminants de son développement. [In French]

Ramirez, O.A., Misra, S., and Field J. (2003). Crop yield distributions revisited. American Journal of Agricultural Economics 85(1): 108-120.

Saha, A., Shumway, C. R., and Talpaz, H. (1994). Joint estimation of risk preference structure and technology using expo-power utility. American Journal of Agricultural Economics 76(2): 173-184.

Siri, A. (2013). Revue diagnostique des dépenses publiques de base dans le secteur agricole (2004–2012) du Burkina Faso : Les dépenses d'intrants. Technical report, Banque mondiale. *[In French]*

Suri, T. (2011). Selection and comparative advantage in technology adoption. Econometrica 79(1): 159-209.

Tack, J., Harri, A., and Coble, K. (2012). More than mean effects: Modeling the effect of climate on the higher order moments of crop yields. American Journal of Agricultural Economics 94(5): 1037-1054.

Tanaka, T., Camerer, C.F., and Nguyen, Q. (2010). Risk and time preferences: Linking experimental and household survey data from Vietnam. American Economic Review 100 (1): 557-571.

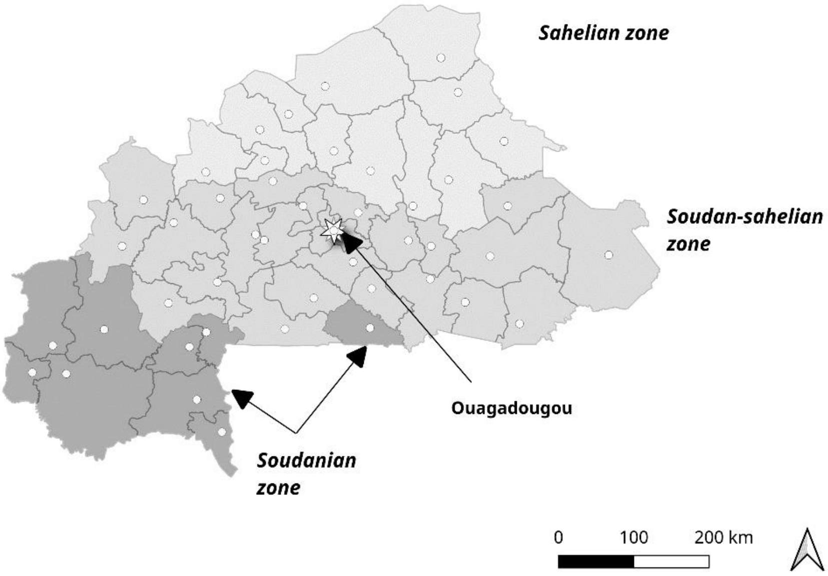
Wanzala-Mlobela, M., Fuentes, P.P., and Mkumbwa, S. (2013). NEPAD agency policy study: practices and policy options for the improved design and implementation of fertilizer subsidy programs in Sub-Saharan Africa. Technical report, NEPAD.

Wik, M., Aragie Kebede, T., Bergland, O., and Holden, S.T. (2004). On the measurement of risk aversion from experimental data. Applied Economics 36(21): 2443–2451.

World Bank Group. (2018). Country partnership framework for Burkina Faso for the period FY18-FY23. World Bank, Washington, DC. World Bank.

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

Appendix A: The three agro-ecological zones



Classification of the provinces in the three agro-ecological zones

Sahelian Zone	Sudan-Sahelian Zone	Sudanese Zone
Oudalan (Gorom-Gorom) Soum (Djibo) Séno (Dori) Yagha (Sebba) Loroum (Titao) Bam (Kongoussi) Sanmatenga (Kaya) Gnagna (Bogandé) Namentenga (Boulsa) Yatenga (Ouahigouya) Zondoma (Gourcy) Sourou (Tougan)	Balé (Boromo) Mouhoun (Dédougou) Tuy (Houndé) Boulgou (Tenkodogo) Tapoa (Diapaga) Sissili (Léo) Banwa (Solenzo) Bazèga (Kombissiri) Boulkiemdé (Koudougou) Ganzourgou (Zorgho) Gourma (Fada N'Gourma) Kadiogo (Ouagadougou) Kompienga (Pama) Kossi (Nouna) Koulpélogo (Ouargaye) Kouritenga (Koupéla) Kourwéogo (Boussé) Nayala (Toma) Oubritenga (Ziniaré) Passoré (Yako) Sanguié (Réo) Ziro (Sapouy) Zoundwéogo (Manga) Komandjoari (Komondjari)	Houet (Bobo-Dioulasso) Comoé (Banfora) Léraba (Sindou) Bougouriba (Diébougou) Poni (Gaoua) Noumbiel (Batié) Kénédougou (Orodara) Ioba (Dano) Nahouri (Pô)

Appendix B - Assessing the presence of selection bias

In order to check if the plots on which some nitrogen has been applied are different from the rest of the plots, we compute the estimated shape parameters setting the quantity of nitrogen at zero and the other covariates at their (sub-)sample mean within each zone. We also report the average estimated mean yield in each zone for plots without any nitrogen.

Table B1. Estimated shape parameters on two different sub-samples, at zone mean and at N=0, with 95% confidence intervals, and estimated average yield by zone

	Sahelian zone	Sudan-Sahelian zone	Sudanese zone
	Conditional Beta distribution estimated on sub-sample of plots *which did not receive any N* (at sample mean & for N = 0)		
α and 95% CI	1.71 [1.52;1.89]	1.63 [1.51;1.74]	1.10 [0.99;1.21]
β and 95% CI	5.14 [4.52;5.75]	3.50 [2.93;4.06]	3.73 [3.21;4.24]
Estimated mean	1,072 kg/ha	1,334 kg/ha	989 kg/ha
# of plots	700	3,050	1,299
	Conditional Beta distribution estimated on sub-sample of plots *on which some N was applied* (at sample mean & for N = 0)		
α and 95% CI	1.06 [0.63;1.49]	1.80 [1.49;2.11]	1.04 [0.74;1.35]
β and 95% CI	3.81 [2.69;4.93]	3.64 [3.06;4.22]	1.95 [1.38;2.52]
Estimated mean	949 kg/ha	1,386 kg/ha	1,454 kg/ha
# of plots	142	1,568	1,042

The upper part of Table B1 shows estimated shape parameters and estimated mean yield for the sub-sample of plots which did not receive any nitrogen. The estimated means for the three agro-ecological zones are found to be pretty close to the observed means as shown in Table 2: the estimated mean yield is 1,072 kg/ha in the Sahelian zone, 989 kg/ha in the Sudanese zone, and 1,334 kg/ha in the Sudan-Sahelian zone, while the actual average yield in the three zones is 1,076 kg/ha, 972 kg/ha and 1,245 kg/ha, respectively. The lower part of Table B1 shows estimates on the sub-sample of plots which received some nitrogen. However, since the expected mean yield is estimated at N=0, these mean estimates cannot be compared to the observed mean yields.

Appendix C: Maize yield distribution with levels of nitrogen at 60 kg/ha and above

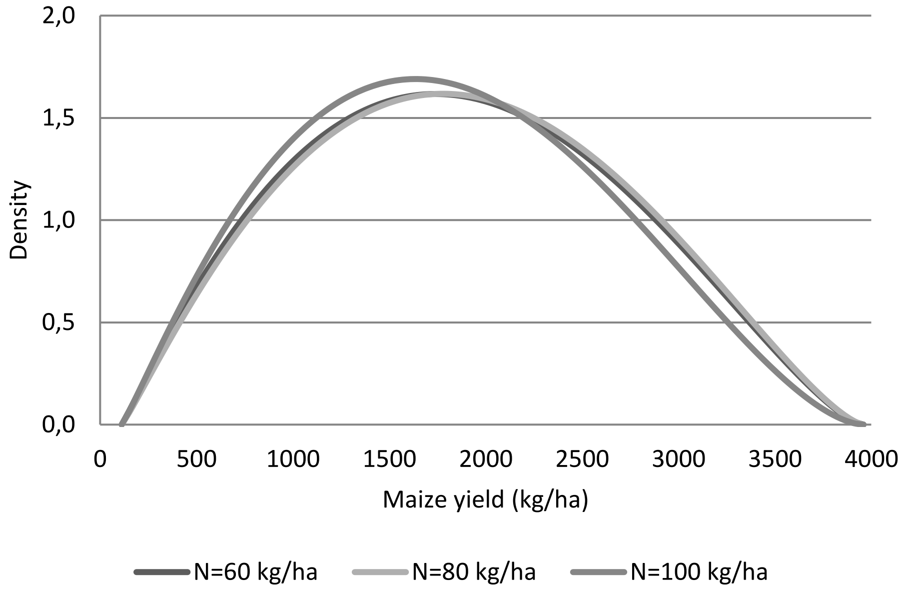


Figure C1. Estimated maize yield distribution on the Sudan-Sahelian zone for levels of nitrogen at 60 kg/ha and above