

Centre for  
Climate Change  
Economics and Policy

An ESRC Research Centre



Grantham Research Institute on  
Climate Change and  
the Environment

# A Simple Framework for the Estimation of Climate Exposure

Xavier Vollenweider

May 2014

Centre for Climate Change Economics and Policy  
Working Paper No. 177

Grantham Research Institute on Climate Change and  
the Environment

Working Paper No. 158

**The Centre for Climate Change Economics and Policy (CCCEP)** was established by the University of Leeds and the London School of Economics and Political Science in 2008 to advance public and private action on climate change through innovative, rigorous research. The Centre is funded by the UK Economic and Social Research Council and has five inter-linked research programmes:

1. Developing climate science and economics
2. Climate change governance for a new global deal
3. Adaptation to climate change and human development
4. Governments, markets and climate change mitigation
5. The Munich Re Programme - Evaluating the economics of climate risks and opportunities in the insurance sector

More information about the Centre for Climate Change Economics and Policy can be found at: <http://www.cccep.ac.uk>.

**The Grantham Research Institute on Climate Change and the Environment** was established by the London School of Economics and Political Science in 2008 to bring together international expertise on economics, finance, geography, the environment, international development and political economy to create a world-leading centre for policy-relevant research and training in climate change and the environment. The Institute is funded by the Grantham Foundation for the Protection of the Environment and the Global Green Growth Institute, and has five research programmes:

1. Global response strategies
2. Green growth
3. Practical aspects of climate policy
4. Adaptation and development
5. Resource security

More information about the Grantham Research Institute on Climate Change and the Environment can be found at: <http://www.lse.ac.uk/grantham>.

This working paper is intended to stimulate discussion within the research community and among users of research, and its content may have been submitted for publication in academic journals. It has been reviewed by at least one internal referee before publication. The views expressed in this paper represent those of the author(s) and do not necessarily represent those of the host institutions or funders.

# A Simple Framework for the Estimation of Climate Exposure

Xavier Vollenweider<sup>1</sup>

## *Abstract*

This article introduces a new methodology to estimate climate exposure at the household's level with the standardized precipitation evapotranspiration index (SPEI) as its building block. As the probability distribution of the SPEI is known, one can easily recover the marginal probability distribution of expected consumption. Furthermore, the approach is simple enough to accommodate quantile regressions and hence offer the opportunity to broaden the scope of the analysis to different categories of the population. I illustrate the methodology with a case study on Ethiopia. I find notably that while poor households in the most remote villages are almost as resilient to a 10-year return period drought as poor households living in the vicinity of a town (up to 20 km), the contrary is true for richer households: the ones living in remote parts of Ethiopia are much more at risk than their suburban counterparts.

---

<sup>1</sup> London School of Economics and Political Science, Department of Geography and Environment, Grantham Research Institute on Climate Change and the Environment, contact: [X.Y.Vollenweider@lse.ac.uk](mailto:X.Y.Vollenweider@lse.ac.uk). I gratefully acknowledge the support from the Wash fellowship and the Rural Economy and Development Programme of Teagasc, the Agriculture and Food Development Authority of the Republic of Ireland. I would like also to extend my thanks to both my supervisors, Salvatore Di Falco and Cathal O'Donoghue.

## I. Introduction

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

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

The emergence of climate change and climate adaptation as a main national and international policy challenge following the Rio Declaration on Environment and Development (1992) has made the estimation of household climate exposure more necessary. Furthermore, anyone can nowadays access daily satellite and weather station precipitation and temperature data over several decades and link them easily

with microeconomic data thanks to GIS software (e.g. Quantum GIS<sup>2</sup>, R<sup>3</sup>). Hence, a new methodology utilizing this climate data bonanza and answering policy needs is required.

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

The methodology proposed in the present article is built on the use of standardized measures of weather. The standardized precipitation index (SPI), first introduced by McKee et al.(McKee et al. 1993, 1995), is a locally and frequency based characterization of precipitation levels. Guttman (Guttman 1998, 1999) widely contributed to its popularisation by showing some of its key advantages over the Palmer Drought Severity Index (Palmer, 1965), the index of choice at the time<sup>4</sup>. The SPI allows the comparison of hydrological conditions across space and time(Hayes et al. 1999), is flexible enough to consider different kind of droughts (e.g. hydrological conditions at months' scale affecting agriculture or at years' scale affecting large-

---

<sup>2</sup> Quantum GIS Development Team (2013). Quantum GIS Geographic Information System. Open Source Geospatial Foundation Project. <http://qgis.osgeo.org>.

<sup>3</sup> R Core Team (2013). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <http://www.R-project.org/>.

<sup>4</sup> The Palmer Drought Severity Index (PDSI) is based on a water balance equation taking into account precipitation, moisture supply, runoff and evaporation demand at the surface level. According to Vicente-Serrano et al. (2010), although some of the weaknesses of the PSDI have been solved by Wells et al. (2004), the main weakness of the PDSI identified by Guttman (1998) has not been addressed: the fixed temporal scale between 9 to 12 months and the fact that PDSI values are affected by conditions up to four years in the past.

scale water management), simple and tractable, and parsimonious in terms of data requirement.

Note that climate change affects both changes in precipitation and temperature. Vincente-Serrano et al. (2010) have proposed the standardized precipitation evapotranspiration index (SPEI) in order to take into account the influence of temperature on hydrological conditions. Its statistical concept and properties are essentially the same as the SPI, although here it is the difference between precipitation and potential evapotranspiration, i.e. the net balance of water, which is standardized. As both temperature and precipitation have an impact on agricultural production and the livelihood of rural populations and as the SPEI is more sensible in the context of Climate change, we settle for the SPEI index as our standardized measure of weather. The use of the SPEI offers the opportunity to easily characterize average production or consumption under locally and frequency-defined weather scenarios. As the framework is very simple, it can easily be extended to quantile regressions in order to broaden the scope of analysis to households at different quantiles of the population distribution. In order to control for unobserved heterogeneity, we rely on penalized quantile fixed effects quantile regressions proposed by Koenker (2004).

Once the climate risk exposure has been estimated, a vulnerability index is needed to summarize the information. We rely on three indices: (1) poverty risk, (2) expected shortfall and (3) relative risk premium. We apply the proposed methodology to the consumption level of rural households in Ethiopia with data from the Ethiopian rural household survey<sup>5</sup>, a panel dataset with seven rounds conducted between 1989 and 2009, including more than 1,200 households. The climate data come from the African Rainfall Climatology Version 2 dataset and the Climate Prediction Center Global Land Surface Air Temperature Analysis

---

<sup>5</sup> The Ethiopian rural household survey data have been made available by the Economics Department, Addis Ababa University, the Centre for the Study of African Economies, University of Oxford and the International Food Policy Research Institute. Funding for data collection was provided by the Economic and Social Research Council (ESRC), the Swedish International Development Agency (SIDA) and the United States Agency for International Development (USAID); the preparation of the public release version of these data was supported, in part, by the World Bank. AAU, CSAE, IFPRI, ESRC, SIDA, USAID and the World Bank are not responsible for any errors in these data or for their use or interpretation.

(GHCN+CAMS) dataset. All datasets used in the present study are freely available online. In Section 2.3, we introduce methodology after a brief discussion of the classical estimation framework; in Section 2.3, the data is presented; in Section 2.4, the results are set out and we conclude in Section 2.5.

## II. Estimation framework

The classical risk estimation methodology was developed when climate and weather data were not widely available. The emphasis was hence on *production risk*, a catch term for drought, flood, pest and animal diseases. In other words, it was viewed as all factors affecting production which are not under the farmer's control, oscillating randomly from year to year and not related to market risk (e.g. inputs and outputs price volatility); resources risk (e.g. fertilizers, seeds and labour supply shocks), institutional risk (e.g. changes in policy), *financial risk* (e.g. changes in the interest rates charged on the debt of the farm), personal risk (e.g. health issues, accidents), and asset risks (thefts or fire damages to buildings, machinery and livestock)(Hardaker et al. 2004; Hazell 1992). Note that financial risk, personal risk, and asset risk are rarely controlled for in applied studies and hence are lumped into production risk.

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

present paper. Its limitation is that it considers the entire sample's rainfall distribution when computing the quantile instead of focusing on localized weather conditions.

The goal of risk estimation has been to estimate the different central moments of the probability distribution of production. The first central moment is the mean, i.e. the expected output or yield. The second moment, i.e. the variance, is a measure of the dispersion of the possible production levels. For instance, a farmer expecting a yield between 200 kg/ha and 4,000 kg/ha would have a higher variance than a farmer expecting a yield between 1,800 kg/ha and 2,200 kg/ha. Variance has hence been one of the first measures of risk. The third moment, summarized by the skewness, is a measure of the asymmetry of possible yields. Negative skewness implies that expected yield is lower than the most likely one and that if bad and good harvests with the same probability are compared, the bad harvest will cost more than the bumpy one could have yield. It is hence often interpreted as a measure of downside risk. The fourth moment, summarized by the kurtosis, is a measure of the *peakedness* of the distribution. For a mono-modal distribution, a high kurtosis implies that most yield levels away from the mode are almost equally likely. This indecision between probabilities is close to the original definition of Knightian uncertainties: a situation in which the agent cannot assign probabilities to the set of possible events.

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

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

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

$$\begin{aligned} E(y) &= E[f(\mathbf{x}, \boldsymbol{\beta}) + h(\mathbf{x}, \boldsymbol{\gamma})^{1/2} \varepsilon] = E[f(\mathbf{x}, \boldsymbol{\beta})] + E[h(\mathbf{x}, \boldsymbol{\gamma})^{1/2}] E[\varepsilon] = E[f(\mathbf{x}, \boldsymbol{\beta})] \\ &= f(\mathbf{x}, \boldsymbol{\beta}) \end{aligned} \quad !)$$

Therefore,  $f(\mathbf{x}, \boldsymbol{\beta})$  is the mean production function, introduced above as the deterministic part of production. As the error term,  $\varepsilon$ , is *iid*, the variance of the production function can be calculated as:



$$\begin{aligned}
\text{var}(y) &= E[y - E(y)]^2 = E \left[ f(x, \beta) + h(x, \gamma)^{\frac{1}{2}} \varepsilon - E \left[ f(x, \beta) + h(x, \gamma)^{\frac{1}{2}} \varepsilon \right] \right]^2 \quad 3) \\
&= E \left[ f(x, \beta) + h(x, \gamma)^{\frac{1}{2}} \varepsilon - E[f(x, \beta)] \right]^2 \\
&= E \left[ h(x, \gamma)^{\frac{1}{2}} \varepsilon \right]^2 \\
&= h(x, \gamma)
\end{aligned}$$

and the marginal impact of a given input  $x$  on variance is:

$$\frac{\partial h(x, \gamma)}{\partial x} = \gamma \quad (4)$$

Input  $x$  can be risk-increasing ( $\gamma > 0$ ), risk-decreasing ( $\gamma < 0$ ) or risk-neutral ( $\gamma = 0$ ). Chavas and Di Falco (2006) extended the model to the third moment, allowing the computation of the marginal effects of inputs on skewness, i.e. on downside risk. Production might hence not only exhibit conditional heteroskedasticity, but also conditional *heteroskewness* and, more generally, all moments can be function of the inputs.

Antle (1983) showed that although input effects on variance are not determined by their effects on the mean, the Just and Pope approach restricts the effect of inputs across variance and higher moments. He proposed the so-called ‘moments based approach’ where the central probability moments (i.e. mean, variance etc.) are directly specified:

$$\mu_1(x, \beta_1) = \int y f(y|x) dy \quad (5)$$

$$\mu_t(x, \beta_t) = \int (y - \mu_1)^t f(y|x) dy \quad t \geq 2 \quad (6)$$

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

mean, variance and skewness of the conditional distribution of each farmer's production.

A limitation of these approaches is that they are highly parametric. Indeed, specification errors in the first moment, respectively Equations (1) and (5), cascade across the whole model, directly affecting the estimation of the higher moments. A popular solution is to choose a flexible functional form such as the translog function, which corresponds to a second order Taylor approximation around the mean of the true production function (e.g. Greene 2003). Although mathematically appealing, the translog functions are notoriously hard to estimate with a sample of a few hundred observations (the usual sample size of rural household surveys): the set of covariates enters the function multiple times - in level, square and through the series of interaction terms - giving rise to important multicollinearity issues<sup>6</sup>. It is hence difficult to obtain statistically significant estimates and no test provides an objective criterion to select which covariates to retain. Full information maximum likelihood estimation and general method of moments provide more efficient results, although issues persist. As Kumbhakar and Tveteras (2003) note: “[t]he idea of dropping insignificant variables is not pursued [...] due to several problems. First, it destroys the flexibility of the mean output function. Second, dropping one insignificant variable caused other insignificant (significant) variables to be significant (insignificant) due to high multicollinearity (which is always present in flexible functions) and the use of a system approach. Furthermore, we found no natural order to select variables for exclusion in the present model”. Therefore, although the conditional expectation might fit well on average, marginal effects are difficult to ascertain.

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

### ***Climate exposure as the marginal distribution of consumption on SPEI***

---

<sup>6</sup> For instance, a production function with four explanatory variables, say labour, fertilizer, land and capital, implies fourteen parameters to estimate.

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

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

Recently, Vicente-Serrano et al. (2010) have proposed focusing on the evapotranspiration index (SPEI) in order to take Climate change into account. The intuition is the same, the only change being that it is not precipitation, but the evapotranspiration index that is standardized. There are several evapotranspiration indices, but they are all based on the same logic, namely the difference between precipitation and potential evapotranspiration. The simplest potential evapotranspiration index in terms of data requirement is the Thornton index. The only data needed are the latitude and the temperature. The detailed derivation of the SPEI can be found in Vicente-Serrano et al. (2010).

The great advantage of using the SPEI (or the SPI) is knowing that it is distributed as a standard normal distribution. Hence, once we have estimated the conditional expectation of production as a function of SPEI, we can easily reconstruct the marginal distribution of the expected production of the average household. We call

this marginal distribution the average climate exposure in contrast to the short-term, or running-season, production risk exposure estimated in the classical framework. We can safely interpret it as the climate exposure given the time scale on which the SPEI is calculated.

Let's formalize the argument. The first step is to estimate the conditional expectation of production on the SPEI variable with an ordinary least square (OLS) regression:

$$y = g(S, \beta) + \varepsilon$$

where  $y$  is production,  $S$  is the SPEI,  $g(S, \beta)$  is the expectation of  $y$  conditional on  $S$  values and parameterized by  $\beta$ , and  $\varepsilon$  is an identically and independently distributed error term. In order to recover the marginal probability density function of  $g(S, \beta)$ , we simply need to compute the marginal distribution of  $S$ , which is known since  $S$  follows a standard normal distribution. Hence, we can compute the marginal density of  $g(S, \beta)$ :

$$f(\mu) = \int_{-\infty}^{\infty} g(S, \beta) f(S) dS$$

where  $f(S)$  is the marginal density of  $S$ :

$$f(S) = \frac{(1/2\pi)e^{-S^2/2}}{\int_{-\infty}^{\infty} (1/2\pi)e^{-S^2/2} dS}$$

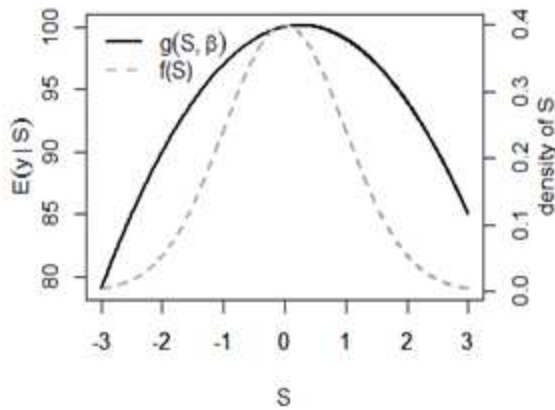
Production is hence characterized under a complete set of local weather scenarios, e.g. from normal conditions to droughts and floods with 100 years' return periods. The marginal distribution,  $f(\mu)$ , can be summarized in different ways. We can simply provide the estimate of its mean, variance, skewness or kurtosis or compute the three indices mentioned in the introduction: poverty risk, expected shortfall or relative risk premium.

Let us illustrate the idea with a simulation exercise. We assume that production reaches a maximum in conditions slightly moister than normal and decreases both with positive and negative values so that the conditional mean has a quadratic form:

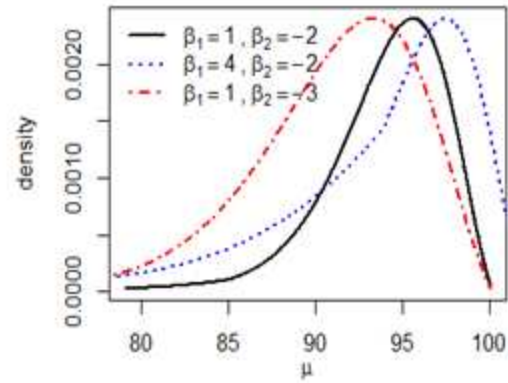
$$g(S, \beta) = \beta_0 + \beta_1 S + \beta_2 S^2$$

Letting  $\beta_0 = 10$ ,  $\beta_1 = 1$  and  $\beta_2 = -2$ , we compute  $g(S, \beta)$  according to a SPEI sequence of 1,000 equally-spaced values on the range  $[-3, 3]$  (Figure 1 (a)). We superimpose in grey the density function of the SPEI (right axis). We then compute the marginal distribution of  $g(S, \beta)$  and plot it on Figure 1 (b) with the plain black line.

**Fig 1a: Conditional Expectation**



**Fig 1b: Climate Exposure**



The intercept parameter has only a location shift effect on the distribution: increasing  $\beta_0$  would increase expected consumption vertically on the Figure 1 (a) and shift  $f(\mu)$  to the right on Figure 2.1 (b) (not represented). By contrast,  $\beta_1$  and  $\beta_2$  have higher order effects. For a given  $\beta_2$ , the greater is  $\beta_1$ , the greater is the mode of  $f(\mu)$ . However, it comes at the cost of greater downside risk as represented by the blue dotted line crossing above the plain black line at 90. For a given level of  $\beta_1$ , an increase in  $|\beta_2|$ , implies a greater variance and more downside risk. We illustrate it by setting  $\beta_2 = -3$  (instead of -2) and drawing on Figure 1 (b) the marginal density of  $\mu$  in red. To sum-up, the parameters  $\beta_1$  and  $\beta_2$  characterize the climate sensitivity of the average household, while the marginal distribution  $f(\mu)$  gives the climate exposure of the average household. We can summarize climate exposure in different vulnerability indices.

Poverty risk is easily obtained by deriving from  $f(\mu)$  the probability of falling below the poverty line. The expected shortfall, also known as the conditional values-at-risk, is a standard risk metric in the finance literature (e.g. Engel and Manganeli, 1999). It is obtained by defining first a probabilistic threshold, for instance a bad event with a probability of occurrence of 5%, and then by computing the difference between the mean under the threshold and the threshold. We modify it slightly by focusing on the shortfall between the poverty line and the lower partial mean under the chosen probabilistically defined threshold:

$$ES(c, K) = c - \int_{-\infty}^K p(S)g(S, \beta) dS$$

where  $c$  is the poverty line and  $K$  is defined in terms of the return period of the event of interest. For instance, for a 35 years return period drought,  $K$  equals -1.9 and the

expected shortfall measures the average cost of bringing back a household to the poverty line in the case of a 35 years return period drought.

Lastly, we can compute the relative risk premium with (e.g. Antle 1987; Chavas and Holt 1996):

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

where  $\mu_i$  is the  $i^{\text{th}}$  central moment, AP is the coefficient of absolute risk aversion (Pratt) for mean-preserving spread aversion, DS is the coefficient of downside risk aversion (Menezes et al. 1980), for mean-spread-preserving skewness preferences and FT is the coefficient of kurtosis aversion (Rubinstein et al. 2006) for mean-spread-skewness preserving kurtosis aversion. We specify the utility function as follows:

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

where  $\gamma$  is the coefficient of relative risk aversion and is set equal to 2 (Ligon and Schechter 2003).

These three vulnerability indices give different perspectives on the climate exposure of households. The poverty risk is probably the best index for evaluating long term development needs as it is mostly affected by the location of the marginal distribution of consumption. For instance, a policy maker interested in having the greatest impact on average poverty should look at the poverty risk indicator and target the population category with the highest poverty risk. The expected shortfall index captures downside risk and is likely to be the most useful for contingency planning, e.g. for the management of emergency food stocks by humanitarian organisations or for designing a safety net programme. Lastly, the relative risk premium emphasizes the trade-off between expected profit and risk and could be used for targeting the roll-out of private agricultural insurance policies such as weather index insurance. Indeed, the relative risk premium, also known as the implicit cost of risk bearing, is an estimate of household willingness to pay for risk reduction.

Note that this simple framework can be extended in several directions. First, we can include other control variables such as inputs and regional dummies in order to estimate how climate exposure varies according to input mixes and regional specificities. In the example above, this would correspond to shifting the parameter  $\beta_0$ . We can also interact these variables with SPEI to investigate the presence of higher order effects. For instance, the interaction term between  $S^2$  and an input

deemed to make farmers more climate resilient should be positive, i.e. should decrease climate exposure.

Second, instead of focusing on the marginal distribution of expected production,  $\mu$ , we can look at the marginal distribution of  $y$  at other quantiles of the  $y$  sample distribution. It is likely that poorer farmers might be more exposed to climate because of a lack of *ex-ante* and *ex-post* risk-mitigating options such as irrigated plots, liquid assets (e.g. bullocks and gold ornaments), off-farm jobs, savings and affluent social networks (e.g. relatives working in the nearby big town). We can therefore expand the analysis from the climate exposure of the average household to the climate exposure at different quantiles:

$$f(\mu_\tau) = \int_{-\infty}^{\infty} g(S, \hat{\beta}_\tau) f(S) dS$$

where  $\mu_\tau = Q_\tau(c_{i,t} | S)$  is the conditional quantile of consumption as a function of SPEI. Panel econometrics methods for quantile regression have been developed by Koenker (2004) and Abrevia and Dahl (2008). They have recently been applied by Bache et. al (2013) to the impact of prenatal maternal smoking on the dispersion of birthweights and by Dahl et al. (2013) to the impact of the decentralization of wage bargaining on wage dispersion. As in the classical mean regression panel methods, they allow for the control of unobserved heterogeneity within the sample.

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

A second version of the paper, with exactly the same set of data, econometric analysis, results and published by the same authors one year later in *Labour Economics*, is entitled “Does education reduce wage inequality?”. In the latter paper, the authors give the inequality interpretation of quantile regressions, i.e. a positive difference between higher and lower quantiles estimates implies that education increases inequality: their “findings imply that schooling may have a positive impact upon within-group wage inequality, as the spread of returns increases for higher educational levels”(Martins and Pereira 2004). The rationale behind this is that “the earnings increment associated to schooling is higher for those individuals whose unobservable characteristics place them at the top of the conditional wage distribution”.

It is hence akin to the latent effect interpretation of quantile regression: inequality in conditional wage outcomes is the result of differences in innate ability revealed by quantile regressions (Koenker 2005). Note that this interpretation is, in turn (and quite paradoxically), related to a special case of Kanbur’s model (1979) where risk is represented by the ability risk that an entrepreneur faces when starting a business for the first time, i.e. the uncertainty about his own capacity to run it. Other earlier works (e.g. Friedman 1953) have drawn the link between risk and inequality. It also echoes the concept of ‘veil of ignorance’ used in thought experiments by political philosophers to apprehend social contracts and redistribution (e.g. Rawls 1971).

We propose to cut the Gordian knot by defining inequality as the between-sample variation captured by quantile regression of consumption and by defining risk as the marginal distribution of consumption before by exploiting the properties of the SPEI. Climate exposure will be summarized in three indices: poverty risk, expected shortfall and the risk premium.

### **III. Data**

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



There are large seasonal fluctuations in consumption as documented by Dercon and Krishnan (2000). As the surveys haven't been conducted exactly at the same period of the year over the rounds, we follow Dercon et al. (2012) and drop data from rounds 2, 3 and 4. Indeed, rounds 2 and 4's data were collected in most villages just after harvest, when a household's consumption is expected to be at its maximum. Round 3 is removed in order to have an equally spaced panel (1994, 1999, 2004, 2009) and avoid hence inconsistent estimates due to heterogeneous frequency (Dercon et al., 2012). The other rounds have been performed, on average, 6 to 9 months after harvest.

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

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

We used two sets of data for the computation of the standardized precipitation evapotranspiration index (SPEI) thanks to the R package SPEI (Beguería and Vicente-Serrano, 2013) with the Thornton evapotranspiration index. The precipitation data come from the African Rainfall Climatology Version 2 dataset (ARC2, Novella and Thiaw 2012), providing daily estimates at a resolution of 0.1 decimal degree from 1983 to the present, and are based on a combination of gauge and satellite data. The dataset has been developed as a key input of the Famine Early Warning System Network (FEWSNET), one of the main indicators used by international humanitarian agencies to monitor food security. The temperature data comes from the Climate Prediction Center Global Land Surface Air Temperature Analysis (GHCN+CAMS, NOAA 2001). They come as monthly mean surface air temperatures at a 0.5 decimal degree resolution over the period from 1948 to the present. One of its recommended uses is precisely the computation of evapotranspiration indices. Both ARC2 and GHCN+CAMS datasets are matched with the ERHS thanks to ward level (*kebele*) administrative boundaries shapefile (Ethiopian Statistical Agency, 2007 census).

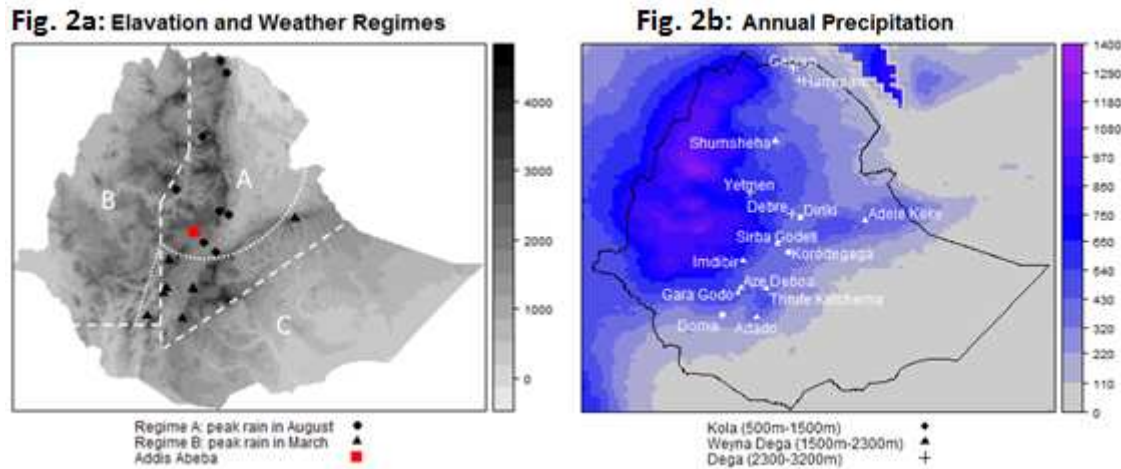
The *kebele*, or Peasant Associations (PA) in the rural part of the countries, were founded by the Coordinating Committee of the Armed Forces, Police, and Territorial Army of Ethiopia, also known as the Derg, after the fall of Emperor Haile Selassie in 1974. They are the lowest administrative unit. We have chosen as matching coordinates the centre of each PA computed with centroids of Voronoi. Note that the median area of the EHRS PAs is smaller ( $50 \text{ km}^2$ ) than the median ARC2 and GHCN+CAMS cells ( $120 \text{ km}^2$  and  $3,025 \text{ km}^2$  on average respectively); they hence constitute a matching metric precise enough for the climate data resolution<sup>7</sup>.

There are three main weather regimes in Ethiopia: the northern part has a bi-modal regime with a long rainy season from June to September and a short rainy season from March to May (regime A); the western part of the country has a mono-modal regime with rainfall from June to September (regime B); and the southern and eastern part has a mono-modal weather regime with rains from February to May (regime C) (NMSA 1996, cited in Abebe, 2010). The approximate hand-drawn partition of the country between weather regimes, according to a map of the Ethiopian National Meteorological Agency (1996) reproduced in Abebe (2010), is mapped with long

---

<sup>7</sup> Area weighted precipitation and temperature means would also have been an option for PAs at the junction of multiple cells, but given the spatial definition of the climate datasets, it would not have affected the results much.

dashed lines in Figure 2 (a). Note that according to the ARC2 rainfall data for each PA, the partition is slightly different (dotted line)<sup>8</sup>.



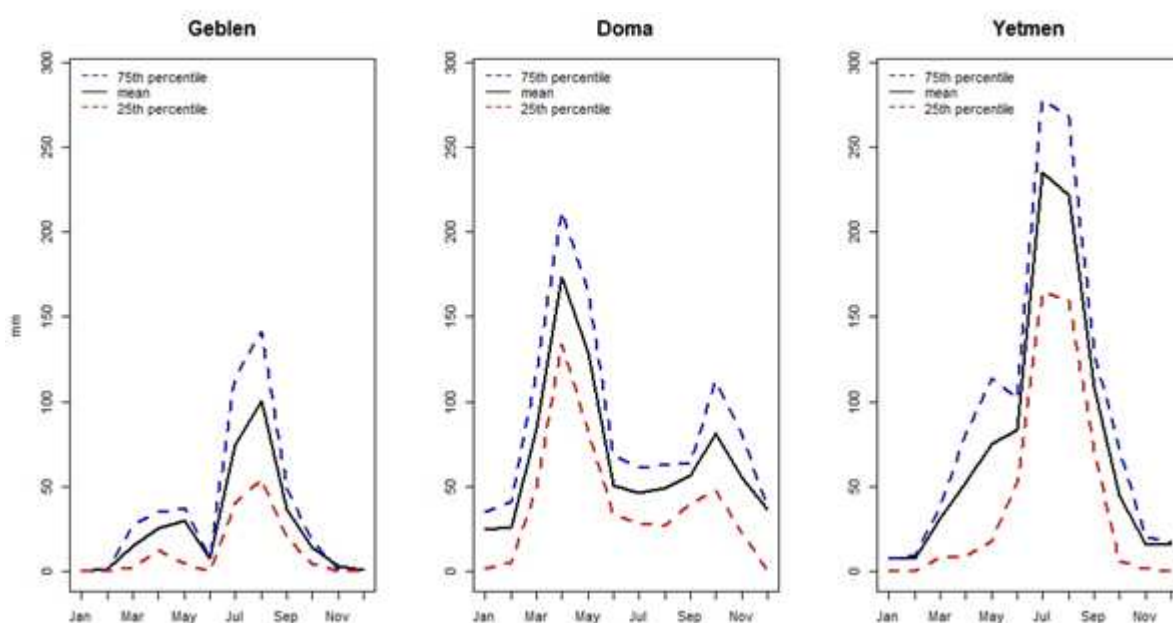
Figures 2: On the left, the long dashed lines are the approximate partition of the country between weather regimes according to a 1996 map of reproduced in Dawit Abebe (2010) and the dotted lines represent an alternative partition matching the ARC2 data at Pas' locations. The map on the right is the average annual precipitation over the period 1990 to 2013. Although it is clear that precipitation concentrated on reliefs because of convective rain, there are great differences in precipitation between PAs located at similar altitudes: Gebelen receives less than 320mm on average while Yetmen, in the same agro-ecological zone, receives twice as much.

Note that the cumulative level of rainfall varies a lot between PAs in regime A (fig. 2b): normal annual precipitation<sup>9</sup> for Gebelen and Harresaw (Tigray region, top North) is only 270 mm while it is 680 mm in Yetmen (Amhara, central North). The PAs located in weather regime C have a maximum amount of cumulative rainfall in March while those located in weather regime A have their maximum in August. We plot in Figure 3 the annual precipitation profile for Gebelen (regime A), Doma (regime C) and Yetmen (regime A). We use the climate data for the peak months in the analysis.

Figure 1: Monthly Precipitation

<sup>8</sup> Although the ARC2 dataset would allow estimating the boundaries between weather regimes with more precision, it is outside of the scope of the present paper.

<sup>9</sup> Normal computed on 1994 to 2013, Hoefsloot 2013, LEAP software.



We use as our dependent variable real consumption per capita as provided in the ERHS. The explicative variables are the 3 months smoothed SPEI at peak rainfall month, the agro-ecological zones, the quality of the road leading to the next town, the distance to the nearest bank, the number of extension agents within the PA and the presence of a non-governmental organisation (NGO) in the PA. Summary statistics are presented in Table 1.

Table1: Summary Statistics

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

Although the national figures brush a rather positive picture for recent years, micro level evidence from the ERHS warrants some caution. While the poverty rate hovers between 45% and 50% until 1995 in the ERHS sample, it decreases to 30% in the next 3 rounds (1997, 1999, 2004) before rising again, above 50% in 2009 (Dercon et al 2012). The average consumption is 78 birr per month (*circa* USD 18) if one focuses

on the 1994, 1999, 2004 and 2009 rounds. There are some substantial variations across years: the 1989, 1994 and 1995 average consumption is around 70 birr; the 1997, 1999 and 2004 average consumption increases to 90 birr while 2009 sees a 34% drop in consumption to 60 birr per month. Consumption per capita includes household-produced food and hence is directly impacted by weather conditions. Details of the real consumption per capita calculation can be found in Dercon and Krishnan (1998). We follow Dercon and Krishnan (1998) in setting the poverty line at the income level required to buy 2,400 calories per day, i.e. 50 birr. The vulnerability indices are hence linked to climate related food insecurity.

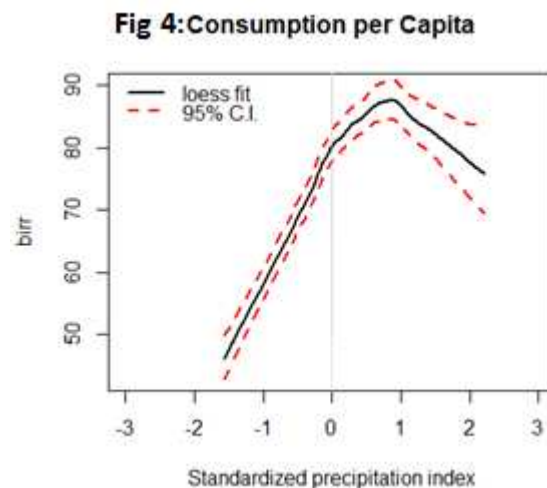
According to the weather regimes identified above, we focus on precipitation in the months of March and August for villages in weather regimes C and A respectively. As we are interested in the hydrological conditions affecting agriculture production, we select the three month smoothed SPEI values. We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

The community-level data capture some of the classical development policies. Indeed, road improvement allows better market linkages with the rest of the country and hence offers better marketing opportunities, larger and more stable sets of products for buying, better price smoothing when local production is adversely hit and allows households to enter into new profitable activities (Dercon, 2012). Extension agents remain a key development mechanism whereby civil servants are dispatched among rural communities to offer farm management advice and increase the adoption of best farming practices. We express it as a dummy variable equal to one if there is at least one extension agent in the PA. Over time, all PAs got an extension agent. The distance to the nearest bank is also of interest as they are a key channel in providing a saving mechanism, as *ex-ante* risk management and credit for adopting more capital

intensive inputs. Furthermore, the distance to the nearest bank serves as a proxy of the *remoteness* or *secludedness* of a particular PA as banks are likely to establish branches in local economic centres. We express it as dummy equal to one if the PA is located at more than 22 km from any bank, the latter value being the median sample distance. The presence of an NGO or a development agency might not only have an impact on their sectorial activity, be it education, health or micro-credit; but they can also be an important provider of jobs for the local community. Furthermore, in case of an adverse climatic shock, an NGO might be able to scale up its activity and to act as a safety net for the local community. Dercon and Krishnan (2003) showed that food aid provided an insurance mechanism.

#### IV. Results

We start by investigating the functional shape of the relationship between real consumption per capita and the standardized precipitation evapotranspiration index with localized polynomial regressions. The smoothing fit is plotted in Figure 4 along the 95% confidence intervals computed by performing 1,000 bootstraps with replacements. The relationship is clearly u-shaped, with a maximum at 0.8, i.e. conditions slightly wetter than normal.



We start hence the analysis with a simple pooled OLS quadratic regression of consumption on SPEI in order to get an idea of the average climate exposure in the sample:

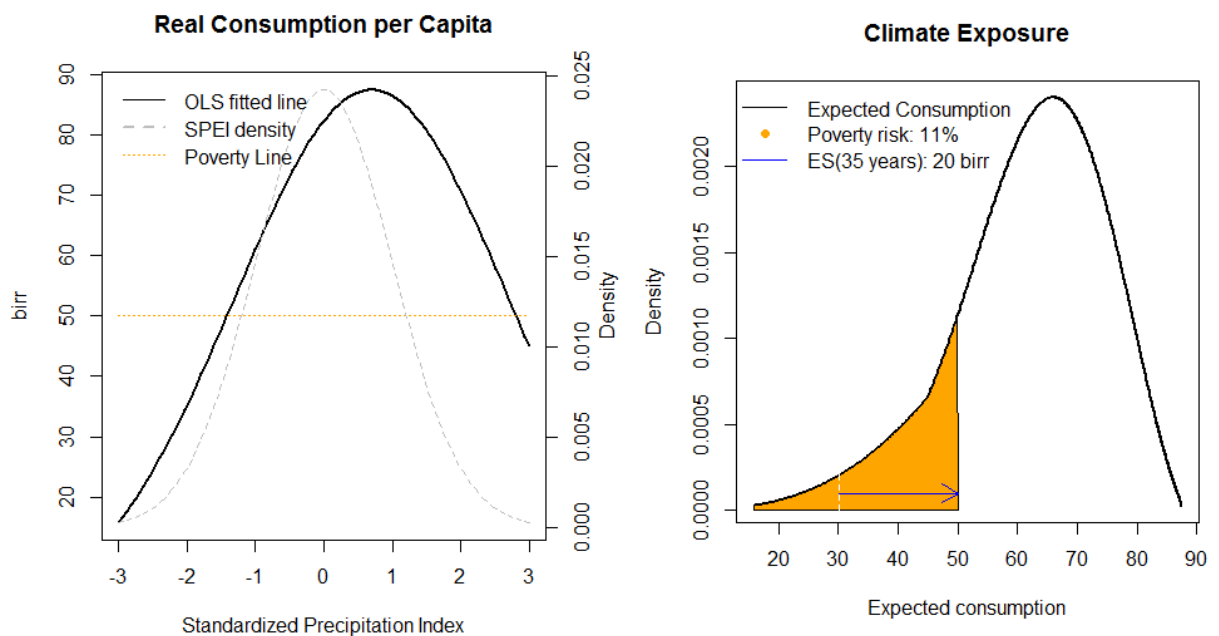
$$\log(c_{it}) = \beta_0 + \beta_1 S_{it} + \beta_2 S_{it}^2 + \varepsilon_{it}$$

where  $c_{it}$  is the real consumption per capita of household  $i$  at time  $t$ ,  $S$  is the 3-months smoothed SPEI at peak rainfall months and  $\beta_j, j = 1,2,3$ , are parameters to be

estimated. Note that the intercept,  $\beta_0$ , is the expected consumption under normal conditions, i.e. when the SPEI equals 0. As the consumption values are very skewed, we apply logarithmic transformation on consumption and compute robust standard errors. Results are presented in Table 2, column 1.

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

Figure 1: Real Consumption per Capita (a) and Climate Exposure (b)



The graph in Figure 5 (a) is the fitted consumption line as a function of SPEI. The probability function of the SPEI is superimposed in grey in order to get a better sense of the likelihood of each SPEI value. The area coloured in orange in Figure 5 (b) is

the probability mass of falling below the hunger line, i.e. 11% in the present case. We also represent the expected shortfall with a 35 years return-period drought (blue arrow, 20 birr). A ‘back of the envelope’ calculation indicates that a 10 years return period drought hitting a region with 100,000 inhabitants would cost a humanitarian agency on average 800,000 birr (*circa* USD 192,000) per month in cash vouchers/transfers to ensure that the basic food requirements are met.

**Table 1: Agro-ecological Zones**

<b>Log (consumption per capita)</b>	<b>Pooled OLS I</b>	<b>Fixed effects OLS</b>	<b>FE QR <math>\tau=0.25</math></b>	<b>FE QR <math>\tau=0.5</math></b>	<b>FE QR <math>\tau=0.75</math></b>
<b>SPEI</b>	0.17*** (0.01)	0.13*** (0.01)	0.12*** (0.02)	0.14*** (0.02)	0.15*** (0.02)
<b>SPEI<sup>2</sup></b>	-0.12*** (0.01)	-0.05*** (0.01)	- 0.08*** (0.02)	-0.05** (0.02)	-0.05** (0.02)
<b>High altitude</b>			0.05 (0.03)	0.09** (0.03)	0.13*** (0.04)
<b>Low altitude</b>			- 0.21*** (0.04)	- 0.11*** (0.03)	-0.08+ (0.04)
<b>High altitude*SPEI</b>		0.09** (0.03)	0.21*** (0.04)	0.15*** (0.03)	0.11** (0.04)
<b>Low altitude*SPEI</b>		0.02 (0.05)	-0.05 (0.06)	-0.08* (0.04)	-0.05 (0.04)
<b>High altitude*SPEI<sup>2</sup></b>		0.05+ (0.03)	-0.01 (0.03)	-0.02 (0.03)	-0.03 (0.03)
<b>Low altitude*SPEI<sup>2</sup></b>		-0.08* (0.04)	-0.01 (0.05)	-0.1*** (0.03)	- 0.15*** (0.04)
<b>Intercept</b>	4.1*** (0.01)		3.8*** (0.02)	4.07*** (0.02)	4.39*** (0.03)
<b>Adjusted R<sup>2</sup></b>	0.04	0.41			
<b>F-test</b>	<0.001	<0.001			

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

$$\log(c_{it}) = \beta_0 + \beta_1 S_{it} + \beta_2 S_{it}^2 + \beta_3 S_{it} * K + \beta_4 S_{it}^2 * K + \beta_5 S_{it} * D + \beta_6 S_{it}^2 * D + \varepsilon_{it}$$

where **K** stands for the lowlands dummy and **D** for the highlands dummy. We test for the presence of unobserved heterogeneity with a Lagrange multiplier test (Breusch-Pagan), an F-test of the model with fixed effects and against pooled OLS (p-value



<0.001) and the Wooldridge test (2002). The null hypothesis is rejected in all cases with a high confidence level (more than 99.99%); we hence conclude that there are important unobserved effects. We then compare the random effects model against fixed effects models with a Hausman test and reject the null hypothesis of convergent estimates, preferring the fixed (*within*) effects model. Lastly, we test for the presence of serial correlation threatening the strict exogeneity assumption of the fixed effects model with the Wooldridge test for serial correlation and fail to reject the null hypothesis of no serial correlation (p-value=0.32). We choose, therefore, a fixed effects model to take into account the households' unobserved heterogeneity. The results are reported in Table 2, column 2.

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

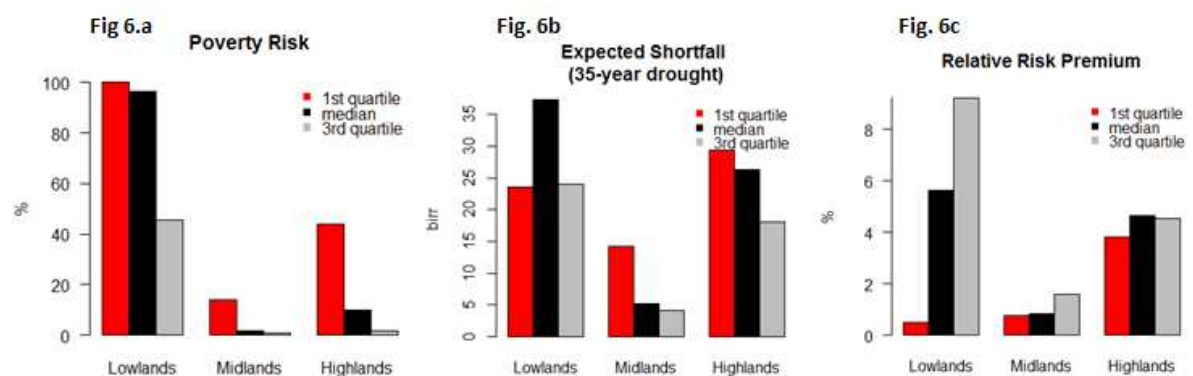
Computing the different indices for each region, the mid-altitude villages have, on average, a poverty risk of 1%, the highlands of 12% and the lowlands of 47%. In terms of expected shortfall, the average household in the midlands is found to be fully resilient even when confronted by a 35 years drought.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

We now present the results across a subset of quantiles of the populations estimated penalized quantile fixed effects quantile regressions (Koenker, 2004) and implemented with the package *rqpd* (Koenker and Bache, 2011). The results are reported in Table 2, columns 3 to 5. Climate sensitivity does not vary much between agro-ecological zones for the lower quartile in terms of the curvature of consumption. The only significant parameter among the interactions is the interaction of the SPEI expressed in level with the highlands dummy: poor households in high altitude villages reach a maximum consumption in conditions wetter than the rest of the sample. Comparing the interaction terms between the lowlands dummy and the SPEI<sup>2</sup>, we see that climate sensitivity increases for households as consumption per capita

increases. It suggests, hence, an important trade-off in the lowlands between increase in consumption and decrease in climate sensitivity, the poorer households being stuck in a low risk-low consumption trap, a phenomenon described in the literature on the risk-induced poverty trap.

We present in Figure 6 (a), (b) and (c) the 3 vulnerability indices across quantiles and agro-ecological zones<sup>10</sup>. In the lowlands, the lowest quartile is trapped in poverty as its poverty risk is 100%. Furthermore, the median households also face a risk of poverty close to 100% while the 3<sup>rd</sup> quartile is slightly above 40%. This contrasts with the results found with OLS where the average household had *only* a 47% risk of poverty. Hence, it is likely that the OLS poverty risk estimate was driven downward by the top percentiles of the population. In the midlands and the highlands, the poverty risk is quite low for households above the median although still substantial for the 1<sup>st</sup> quartile.

The results in terms of the expected shortfall are presented in Figure 6 (b). Although the ranking of agro-ecological zones in terms of risk is respected, the differences are much smaller. Furthermore, the ranking within zones changes a lot, e.g. in the lowlands the median 35-year drought expected shortfall is higher than the lower quartile one. The relative risk premium (Figure 6 c), i.e. the implicit cost of risk, confirms the interpretation of a risk-induced poverty trap by showing that poor households have a smaller relative risk exposure, i.e. they have already reduced risk exposure to its maximum at the cost of a decrease in consumption.



A policy maker interested in having the greatest impact on average poverty with, for instance, the provision of subsidized fertilizers, should look at the poverty risk indicator and target the lowlands. Interestingly, the expected shortfall shows that in

<sup>10</sup> Note that the quantile regressions were run in level to compute the indices because it is *a priori* not clear how to deal with the residuals of exponential quantile regressions when computing the conditional quantiles.

the case of a serious drought, it might not be the poorest quartile of the population which will require most help in the lowlands but instead the median households because the latter are more exposed to downside climate shocks. Lastly, the relative risk premium shows that the implicit cost of risk bearing is the highest among richer households, particularly in the lowlands. Hence, the higher quantile of the population manage to get higher consumption at the cost of a large increase in risk and should therefore be willing to swap part of this risk against some kind of consumption insurance, be it index based or of the traditional agricultural kind.

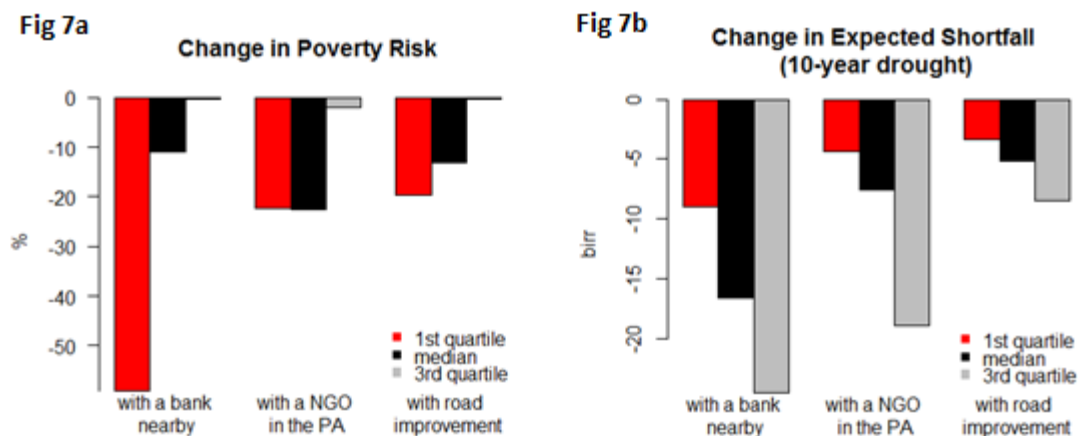
Let's turn now to characteristics which have evolved over time at the community level. The panel is shorter as the community level data are only available for rounds 4, 6 and 7, i.e. 1997, 2004 and 2009. As noted in the data section, the 1997 round was conducted earlier in the season and hence might introduce some unobserved heterogeneity. We attempt to control for it by adding a year dummy for 1997. We focus on the presence of an improvement in the road leading to the next town, the number of extension agents within the PA and the distance to the nearest bank, and the presence of a non-governmental and/or international organization office in the PA. The results are presented in Table 3.

Table 2: Community development factors

Log (consumption per capita)	Fixed effects	FE QR	FE QR	FE QR
	OLS	$\tau=0.25$	$\tau=0.5$	$\tau=0.75$
<b>1997 dummy</b>	0.12 <sup>***</sup> (0.03)	0.1 <sup>+</sup> (0.05)	0.18 <sup>***</sup> (0.04)	0.13 <sup>+</sup> (0.05)
<b>Highlands</b>		0.1 <sup>*</sup> (0.04)	0.05 <sup>+</sup> (0.03)	0.06 (0.04)
<b>Lowlands</b>		-0.06 <sup>+</sup> (0.03)	-0.08 <sup>***</sup> (0.03)	-0.08 <sup>*</sup> (0.03)
<b>Distance to the bank (=1 if &gt; 22 km)</b>	-0.21 <sup>***</sup> (0.04)	-0.27 <sup>***</sup> (0.04)	-0.32 <sup>***</sup> (0.03)	-0.31 <sup>***</sup> (0.04)
<b>NGO office in the PA</b>	0.31 <sup>***</sup> (0.04)	0.11 <sup>*</sup> (0.04)	0.14 <sup>***</sup> (0.04)	0.2 <sup>***</sup> (0.05)
<b>Extension agent</b>	-0.08 <sup>+</sup> (0.04)	0.03 (0.04)	-0.02 (0.04)	-0.04 (0.04)
<b>Road improvement</b>	0.05 <sup>+</sup> (0.03)	0.05 (0.03)	0.1 <sup>**</sup> (0.03)	0.13 <sup>***</sup> (0.04)
<b>SPEI</b>	0.11 <sup>***</sup> (0.02)	0.19 <sup>***</sup> (0.02)	0.13 <sup>***</sup> (0.02)	0.17 <sup>***</sup> (0.02)
<b>SPEI<sup>2</sup></b>	-0.08 <sup>***</sup> (0.02)	-0.12 <sup>***</sup> (0.02)	-0.13 <sup>***</sup> (0.02)	-0.13 <sup>***</sup> (0.02)

Intercept	3.83 <sup>***</sup> (0.05)	4.19 <sup>***</sup> (0.04)	4.52 <sup>***</sup> (0.05)
Adjusted R <sup>2</sup>	0.7		

The 1997 dummy is positive, as expected, because the 1997 round was conducted earlier in the season when consumption is higher. The distance to the bank dummy, equal to one if the PA is located at more than 22 km from any bank, is strongly negative: the average household in such a PA has an expected consumption per capita 21% lower than those in PAs closer to a centre with a more vibrant economy. Note that the effect is quite stable across quantiles of the population (although lower). By contrast, the presence of an NGO office in the PA benefits mostly the median household and above. This might be linked to the fact that jobs created by NGOs tend to benefit the better educated and wealthier households, or it might reveal the difficulty for NGOs to reach the poorest of the poor. Interestingly, road improvement seems again to be of greatest benefit to richer households as no consumption-increasing effect linked to road improvement is found significant in the 1<sup>st</sup> quartile regression. Two specifications were tried for the extension agents: a dummy equal to 1 if there is at least one extension agent (results presented above) and the number of extension agents (results available on request). Although extension agents are found to have a *negative* impact on expected consumption, it is only significant at a p-value of 0.054 and the effect disappears in the quantile regression, so that it is likely to be driven by outliers. Furthermore, once specified in plain numbers, the effect of the extension agents is positive and significant: each additional extension agent increases expected consumption by 10% and has the most impact on the median of the distribution.



We present in Figures 7 (a) and (b) the impact on poverty risk and on the expected shortfall (computed for a 10 years drought) of the variables found significant at different quantiles of the population. As expected, the greatest effect is clearly the bank dummy, capturing the effect of living close to an economic centre (not more than 20 km). This large effect might be due to more off-farm opportunities or to the support from relatives living in these economic centres. Note that, however, in the case of a 10-year drought, poor households living nearby economic centres are not much less exposed than their counterparts in more remote regions, as shown with the expected shortfall. By contrast, richer categories are much more exposed in remote parts compared to their counterparts living close to an economic centre. NGOs and road improvements have a similar effect on the risk of poverty and expected shortfall. Both are, incidentally, positively correlated and it is likely that logistical reasons favour the installation of NGOs in PAs with better road access. Again, we see this mirror relationship between poverty risk and expected shortfall: these are the poorest households who benefit the most in terms of poverty risk reduction but the richest ones in terms of reduction of downside risk.

## **V. Conclusion**

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

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

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

The present paper could be extended in several directions. First, variables on the farm inputs and output mixes could be added in the regressions. Second, the impact of climate adaptation farm strategies could be tested. Lastly, the distributional impact of climate could be better ascertained, either at a micro or macro scale.

### **Acknowledgments**

I thank both my supervisors, Prof. Salvatore Di Falco and Prof. Cathal O'Donoghue, for their excellent guidance during the whole process of the PhD. I would like also to gratefully acknowledge the 4 years' financial support of the Walsh fellowship, Teagasc.

## VI. References

- Abrevaya, Jason and Dahl, Christian M (2008), 'The effects of birth inputs on birthweight: evidence from quantile estimation on panel data', *Journal of Business & Economic Statistics*, 26 (4), 379-97.
- Antle, John M (1983), 'Testing the stochastic structure of production: a flexible moment-based approach', *Journal of Business & Economic Statistics*, 1 (3), 192-201.
- (1987), 'Econometric estimation of producers' risk attitudes', *American Journal of Agricultural Economics*, 69 (3), 509-22.
- Bache, Stefan Holst Milton, Dahl, Christian Møller, and Kristensen, Johannes Tang (2013), 'Headlights on tobacco road to low birthweight outcomes', *Empirical Economics*, 44 (3), 1593-633.
- Chambers, Robert G. and Quiggin, John (1998), 'Cost Functions and Duality for Stochastic Technologies', *American Journal of Agricultural Economics*, 80 (2), 288-95.
- Chavas, Jean-Paul and Holt, Matthew T. (1996), 'Economic Behavior Under Uncertainty: A Joint Analysis of Risk Preferences and Technology', *The Review of Economics and Statistics*, 78 (2), 329-35.
- Dahl, Christian M, Le Maire, Daniel, and Munch, Jakob R (2013), 'Wage dispersion and decentralization of wage bargaining', *Journal of Labor Economics*, 31 (3), 501-33.
- Dercon, Stefan and Krishnan, Pramila (1998), *Changes in poverty in rural Ethiopia 1989-1995: measurement, robustness tests and decomposition* (Centre for the Study of African Economies, Institute of Economics and Statistics, University of Oxford).
- (2000), 'Vulnerability, seasonality and poverty in Ethiopia', *The Journal of Development Studies*, 36 (6), 25-53.
- Dercon, Stefan and Christiaensen, Luc (2011), 'Consumption risk, technology adoption and poverty traps: Evidence from Ethiopia', *Journal of Development Economics*, 96 (2), 159-73.
- Dercon, Stefan, Hoddinott, John, and Woldehanna, Tassew (2012), 'Growth and chronic poverty: Evidence from rural communities in Ethiopia', *Journal of Development Studies*, 48 (2), 238-53.
- Deressa, Temesgen T, Hassan, Rashid M, and Ringler, Claudia (2009), *Assessing household vulnerability to climate change* (Intl Food Policy Res Inst).
- Di Falco, Salvatore and Chavas, Jean-Paul (2006), 'Crop genetic diversity, farm productivity and the management of environmental risk in rainfed agriculture', *European review of agricultural economics*, 33 (3), 289-314.
- (2009), 'On crop biodiversity, risk exposure, and food security in the highlands of Ethiopia', *American Journal of Agricultural Economics*, 91 (3), 599-611.
- Friedman, Milton (1953), 'Choice, chance, and the personal distribution of income', *The Journal of Political Economy*, 277-90.
- Greene, William H (2003), *Econometric analysis* (Pearson Education India).
- Groom, Ben, et al. (2008), 'The story of the moment: risk averse cypriot farmers respond to drought management', *Applied Economics*, 40 (3), 315-26.
- Guttman, Nathaniel B (1998), 'Comparing the palmer drought index and the standardized precipitation index1', *JAWRA Journal of the American Water Resources Association*, 34 (1), 113-21.

- (1999), 'ACCEPTING THE STANDARDIZED PRECIPITATION INDEX: A CALCULATION ALGORITHM1', *JAWRA Journal of the American Water Resources Association*, 35 (2), 311-22.
- Hardaker, J. B., et al. (2004), *Coping With Risk in Agriculture* (2 ed. edn.; Cambridge: CABI).
- Hayes, Michael J, et al. (1999), 'Monitoring the 1996 drought using the standardized precipitation index', *Bulletin of the American Meteorological Society*, 80 (3), 429-38.
- Hazell, Peter BR (1992), 'The appropriate role of agricultural insurance in developing countries', *Journal of International Development*, 4 (6), 567-81.
- Just, Richard E and Pope, Rulon D (1978), 'Stochastic specification of production functions and economic implications', *Journal of Econometrics*, 7 (1), 67-86.
- Kanbur, Steven M (1979), 'Of risk taking and the personal distribution of income', *The Journal of Political Economy*, 769-97.
- Koenker, Roger (2004), 'Quantile regression for longitudinal data', *Journal of Multivariate Analysis*, 91 (1), 74-89.
- (2005), *Quantile regression* (Cambridge university press).
- Koenker, Roger and Bassett, Gilbert (1978), 'Regression quantiles', *Econometrica: journal of the Econometric Society*, 33-50.
- Koundouri, Phoebe, et al. (2009), 'The effects of EU agricultural policy changes on farmers' risk attitudes', *Eur Rev Agric Econ*, jbp003.
- Kumbhakar, Subal C and Tveterås, Ragnar (2003), 'Risk preferences, production risk and firm heterogeneity\*', *The Scandinavian Journal of Economics*, 105 (2), 275-93.
- Ligon, Ethan and Schechter, Laura (2003), 'Measuring Vulnerability\*', *The Economic Journal*, 113 (486), C95-C102.
- Love, H Alan and Buccola, Steven T (1991), 'Joint risk preference-technology estimation with a primal system', *American Journal of Agricultural Economics*, 73 (3), 765-74.
- Martins, Pedro S and Pereira, Pedro T (2004), 'Does education reduce wage inequality? Quantile regression evidence from 16 countries', *Labour Economics*, 11 (3), 355-71.
- McKee, Thomas B, Doesken, Nolan J, and Kleist, John (1993), 'The relationship of drought frequency and duration to time scales', *Proceedings of the 8th Conference on Applied Climatology* (17: American Meteorological Society Boston, MA), 179-83.
- (1995), 'Drought monitoring with multiple time scales', *Ninth Conference on Applied Climatology. American Meteorological Society, Boston*.
- Menezes, Carmen, Geiss, Charles, and Tressler, John (1980), 'Increasing downside risk', *American Economic Review*, 70 (5), 921-32.
- Moss, Charles B (2010), *Risk, uncertainty and the agricultural firm* (World Scientific).
- Rawls, John (1971), 'A theory of justice Oxford University Press', *New York*.
- Rubinstein, Mark, Jurczenko, Emmanuel, and Maillet, Bertrand (2006), *Multi-moment asset allocation and pricing models* (399: John Wiley & Sons).
- Saastamoinen, Antti (2013), 'Heteroscedasticity or Production Risk? A Synthetic View', *Journal of Economic Surveys*.
- Sandmo, Agnar (1971), 'On the Theory of the Competitive Firm Under Price Uncertainty', *The American Economic Review*, 61 (1), 65-73.
- Smale, Melinda, et al. (1998), 'The contribution of genetic resources and diversity to wheat production in the Punjab of Pakistan', *American Journal of Agricultural Economics*, 80 (3), 482-93.



Telhado Pereira, Pedro and Silva Martins, Pedro (2002), 'Is there a return–risk link in education?', *Economics Letters*, 75 (1), 31-37.

Vicente-Serrano, Sergio M, Beguería, Santiago, and López-Moreno, Juan I (2010), 'A multiscalar drought index sensitive to global warming: The standardized precipitation evapotranspiration index', *Journal of Climate*, 23 (7).

Wells, Nathan, Goddard, Steve, and Hayes, Michael J (2004), 'A self-calibrating Palmer drought severity index', *Journal of Climate*, 17 (12).