Essays in Agricultural Economics

Francisco Pereira Fontes

A thesis submitted to the Department of Geography and Environment
of the London School of Economics and Political Science
for the degree of Doctor of Philosophy

February 2018

THE LONDON SCHOOL OF ECONOMICS AND POLITICAL SCIENCE

Declaration

I certify that the thesis presented for examination for the PhD degree of the London School of Economics and Political Science is solely my own work unless otherwise clearly indicated.

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

I declare that my thesis consists of approximately 63,897 words (excluding appendices).

Statement of Conjoint Work

Chapter 3 was co-authored with Dr. Benjamin Groom. My contribution amounted to 70% of the paper.

Chapter 5 was co-authored with Ashley Gorst and Dr. Charles Palmer. My contribution amounted to 50% of the paper.

Chapter 6 was co-authored with Ashley Gorst and Dr. Charles Palmer. My contribution amounted to 50% of the paper.

Abstract

This thesis explores topics in Agricultural Economics and is composed of five papers. In the first paper (Chapter 2), a latent-class stochastic frontier model is used to estimate efficiency scores of farmers in Ethiopia. Compared to conventional models, which assume a unique frontier, much lower inefficiencies are found, suggesting that part of the inefficiencies uncovered in the literature could be an artefact of the methods used. The second paper (Chapter 3) revisits the link between cereal diversity and productivity using a panel dataset in Ethiopia. The results suggest that the positive effect between cereal diversity and productivity becomes much smaller when households who produce teff (a low-productivity and high-value crop) are excluded from the sample, hinting at the possibility that results could be driven by yield differentials between cereals, rather than diversity. The third paper (Chapter 4) estimates the labour impacts of the adoption of Soil and Water Conservation technologies (SWC) in Ethiopia. The results suggest that adopting SWC technologies leads to an increase in adult and child labour. Understanding the labour impacts is important in itself, but it also raises concerns about using impact evaluation methods that require no change in inputs as an identifying assumption of impacts. Paper 4 (Chapter 5), assesses the pertinence of a drought index that has recently been proposed in the literature by Yu and Babcock (2010) and argues that it defines drought too narrowly. An extension to this index is proposed and we show, using a dataset of Indian districts, that the original index is likely to underestimate the impacts of drought. In Paper 5 (Chapter 6), we identify data-driven ranges of rainfall for which the marginal effects of a rainfall-temperature index (RTI) are different and then we discuss how the impacts of drought have changed over the 1966-2009 period in India. Finally, Chapter 7 concludes.

Contents

List of Figures

List of Tables

Acknowledgements

First and foremost, I would like to thank my two supervisors, Dr. Benjamin Groom and Dr. Charles Palmer, for their support, comments, opinions and criticisms throughout my PhD. I would also like to thank my co-author and friend, Ashley Gorst.

I am also grateful to the Economic and Social Research Council for the funding throughout this research, without which this project would have not been possible.

I would also like to thank my parents, Manuela and Hilário, as well as Dalia Fadly for their constant support, which has always meant a lot to me.

Finally, I am also grateful to all those with whom I had the pleasure to work and/or interact during my time as a PhD student, including members of the Department of Geography and the Environment at LSE, members of the Grantham Research Institute on Climate Change and the Environment, and participants at various conferences.

Chapter 1

Introduction

An agricultural sector that is resource-efficient, climate-smart and that enables smallholder farmers to escape poverty will be key to achieving the goals of the international community, enshrined in the Sustainable Development Goals (SDGs). However, anticipated demographic changes and the looming threat of climate change mean that achieving these goals will be very challenging. High-quality, policy-relevant research will be important to inform policy-makers about concrete actions that may deliver positive results. In order for research to be policy-relevant, it is important that it:

- **Uses methods and definitions that are fit-for-purpose**
- **Takes into account the local context**
- **Analyses impacts of policies on a wide-range of metrics**

Failing to do the above could lead, among other things, to estimation biases, misguided policies and unanticipated consequences of policies.

This dissertation unfolds in seven chapters, which focus on Ethiopia (Chapters 2-4) and India (Chapters 5-6). All the chapters look at one or more of the three aspects highlighted above.

In Chapter 2, I focus on the issue of production inefficiencies in Africa and argue that they are partly an artefact of the methodologies commonly used to compute efficiency scores. Conventional econometric methods to compute efficiency scores tend to assume that all the units of observation can be compared to a single production frontier, irrespective of the agro-ecological zone and the technology used. In this chapter, I relax this assumption slightly by allowing more than one frontier. Doing this decreases the potential gains from tackling inefficiencies by more than half.

In Chapter 3, my co-author and I re-examine the link between cereal diversity and agricultural productivity. Similar to previous literature, we find a positive relationship between the two variables. However, upon further exploration, we find that the positive effects are drastically reduced when households who cultivate a low-productivity, high-value crop (teff) are excluded. In our setting, it thus seems that the positive cereal-diversity relationship is partly driven by yield differentials between crops rather than complementarity or a facilitation effect. This also suggest that in situ conservation has potential for development, but only for certain crop mixes.

Chapter 4 examines the impact of adopting Soil and Water Conservation (SWC) technologies on labour outcomes in Ethiopia. I find that adopting SWC technologies leads to an increase in adult and child labour. This is important for two reasons. First, being aware of these impacts is important for policy-makers when deciding what policies to pursue. Second, for researchers focusing on the estimation of the productive impacts of SWC technologies, this paper suggests that using econometric methods that assume no changes in inputs as a result of adoption may be problematic.

In chapter 5, my co-authors and I revisit a drought index that has recently been proposed in the literature by Yu and Babcock (2010). This index is appealing in that it simultaneously incorporates rainfall and temperature in a simple way. However, we argue that it is inadequate in terms of its coverage of dry events, as it excludes all dry (below-normal rainfall) events that occur in cold (below-average temperature) years ("cold droughts"). We therefore propose an alternative index that has all the benefits of the previous index and extends the coverage of events. We then show that ignoring the class of events not included in the original index can lead to a severe underestimation of drought impacts on yields.

In chapter 6, using the index proposed in chapter 5, my co-authors and I focus on two questions. The first question is concerned with the non-linear impacts of extreme rainfall events (droughts and floods) and we use a method (fixed effects threshold model) that allows us to establish data-driven ranges of precipitation where the marginal impacts of our rainfall-temperature index are different. The second question assesses whether, overall, India has become more resilient to drought over time. Concerning the first question, overall, we find that negative impacts exist at low negative deviations from “normal rainfall (above the threshold previously used by the Indian Government to define a drought). We also find that the results differ sharply by agro-ecological zone and crop. Areas that are more arid tend to witness lower impacts at low negative deviations from normal rainfall. However, at negative deviations exceeding a certain threshold, impacts tend to be very large compared to more humid areas. Similarly, crops that are more drought-resilient (millet, sorghum and maize) also tend to have lower impacts at small negative deviations from normal rainfall, but exhibit very large impacts at medium-large negative deviations from normal rainfall. With regards to the results over time, we find that impacts of droughts decreased over time until the 1990s, but this trend has been reversed since the beginning of the millennium. This pattern is broadly consistent across crops (although less strong for rice) and agro-ecological zones. We argue that one possible explanation for the increase in drought impacts since the millennium could be due to a change in rainfall patterns. Specifically, we find that drought-affected districts have seen a decrease in the rainfall received in the year preceding a dry year.

Chapter 7 summarizes the main findings and discusses the policy-implications of this dissertation as well as potential avenues for future research.

Chapter 2

“Blatent” Heterogeneity: Implications for Efficiency Measurement and Policy. A case study of Ethiopia.

Abstract

African Agriculture is often seen as trailing behind the rest of the world and production "inefficiencies", widely documented in the literature, are often seen as a culprit of this under-development. However, most studies assume a common technology across households, thereby ignoring differences in the chosen technology. As a result, it is possible that: 1) Heterogeneity may be mislabelled as inefficiency, with households being compared to a potentially inappropriate frontier; and 2) Different technologies may exhibit very distinct production elasticities.

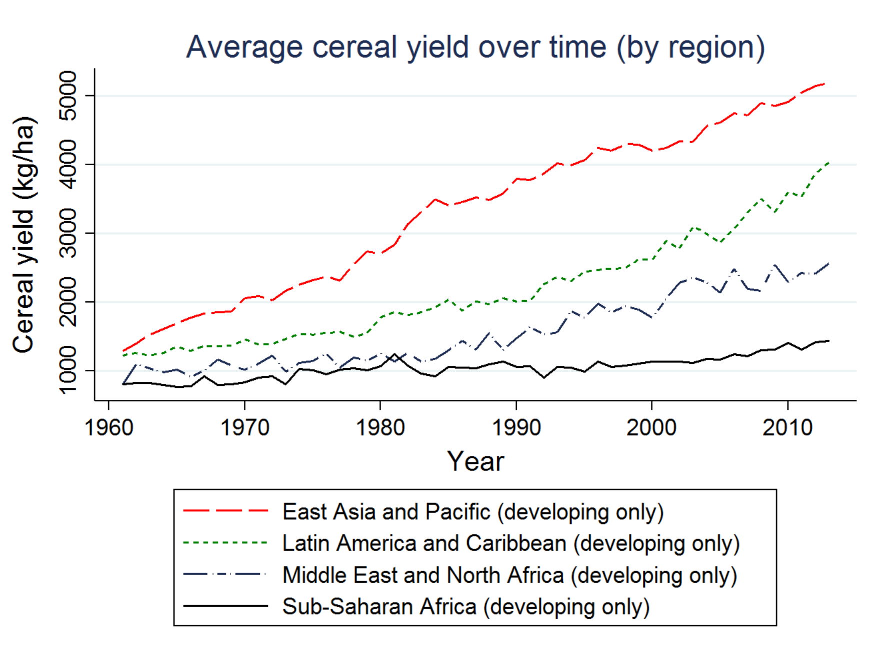
Using data from the Ethiopian Rural Household Survey we use a two-class latent class stochastic frontier model and compare this to conventional stochastic frontier models. We find that: 1) The overall efficiency scores increase from 0.61-0.62 in models that assume a common technology to about 0.79 in the latent class model. This means that the estimated potential gains in cereal production decrease from 61%-64% to about 22%. The results suggest that potential gains from tackling "inefficiencies" in production are lower than previously estimated, questioning the absolute importance of tackling inefficiencies as a policy priority. 2) Large differences in the estimates of the input elasticities of production emerge across latent classes supporting the idea that agricultural policies should account for technological differences.

JEL classification: Q10, Q12, Q18, Q50

2.1 Introduction

Sub-Saharan Africa is often perceived to be lagging behind the rest of the world in terms of its agricultural productivity and the trends in cereal yields in recent decades are often used to highlight this point (Figure 2.1). Production inefficiencies are often seen as a culprit of this underdevelopment and this paper seeks to understand whether part of this belief is an artefact of the methodology used to measure inefficiencies. In particular, it seeks to explore what are the implications of relaxing the underlying assumption of a common technology in terms of the estimated efficiency scores and elasticities.

Figure 2.1: Cereal Yields (kg/ha) by region 1960-2013



Data source: World Bank - World Development Indicators

This seems particularly important given the importance of agriculture in Africa for its development. Collier and Dercon (2014), for example, uphold that, if Africa is ever to experience sustained levels of economic development, the productivity of its agricultural sector will have to change "beyond recognition"; a difficult task, especially in an era marked by unprecedented demographic pressure and increasing environmental concerns.

In such a context, doing more with less would seem like a priority. Thus, addressing the seemingly large production inefficiencies of smallholder farmers in Africa, widely documented in the literature, appears as an obvious policy choice. A meta-analysis of 442 studies of efficiency in African agriculture reveals an average efficiency estimate of 0.68, with this number being lower for cereal crops (Ogundari, 2014). The implications of this are far-reaching, and would imply that, simply by using all inputs "efficiently", an increase in agricultural output in excess of 30 percentage points (or 47%) could be achieved, and this number is even higher for cereal crops, which are often the cornerstone of government agricultural policy.

While studies reviewed in Ogundari (2014) differ widely in terms of both the output analysed and the geographical region of the study, the overwhelming majority tend to assume a common underlying production technology for every unit in the sample or sub-samples of interest. This assumption is, in fact, typical of most models in Stochastic Frontier Analysis (Kumbhakar and Orea, 2004). A case in point are the studies performed at the national level in a number of African countries including Malawi (Tchale, 2009) and Ethiopia (Bachewe et al., 2011, Mekonnen et al., 2013 and Bachewe et al., 2015), where at least one specification assumes a common of the production function across the entire country[1]. This paper argues that doing so has two practical implications of policy relevance.

First, different farming technologies have wildly different agricultural potentials and may require different management practices and inputs. However, most studies still assume a common technology, thus potentially providing a misleading representation of the true potential of a certain technology in a given environment.

Secondly, assuming a common production function across farming systems and regions provides unique average estimates of the output elasticities of inputs. However, being able to account for the heterogeneity in output elasticities of inputs is essential for a successful targeting of agricultural policy. For instance, Pender and Gebremedhin (2008) argue that irrigation is an important complement to successful use of fertilizer. Therefore, moisture strained areas may have a very different (generally lower) output elasticity of fertilizer than those which have abundant irrigation.

The consequences of this are twofold. First, heterogeneity may be mislabelled as inefficiency

[1]In the case of Bachewe et al. 2011, however, in addition to the full sample estimates, the authors also estimate stochastic frontiers for different sub-samples, including fertilizer users and estimates by agro-ecological zones. However, these sub-samples still assume a common production frontier either across fertilizer users across the country or a common production function in a given agro-ecological zone.

(Kumbhakar and Orea, 2004) as households are compared to a frontier that may not be the most appropriate for them. As a result the actual gains from tackling "inefficiencies" in African agriculture may be misstated, and most likely, overstated. Secondly, disregarding heterogeneity in production may, ironically, be an inefficient way of spurring efficiency. The needs of farm households are likely to be diverse, given their natural environment and their farming systems. However, the differentiated needs of farmers contrast sharply with the unique set of elasticities often assumed in the literature.

This paper seeks to illustrate the effects of relaxing the assumption of a common production function on efficiency measurement and elasticities. An application to Ethiopia will be used and it seems pertinent for a number of reasons. First, Ethiopia has a high degree of heterogeneity in both agro-ecological zones and production systems. Second, the Ethiopian Government has identified agriculture as a pillar (Alemu et al., 2009) of its national development strategy, implying this is a topic of great policy relevance. Finally, a wide range of studies exist on Ethiopia and indicate a wide range of efficiency scores (0.37-0.86). Studies at the regional scale tend to indicate substantially higher efficiency scores (0.65-0.86) compared to studies performed at the national level (0.37-0.76). These results are material since more localized studies in Ethiopia indicate a potential efficiency gains in output in the 16%-53% range, as opposed to the 33%-270% (66%-117% in papers using the same dataset) range suggested in the national studies. This discrepancy hints at the possibility that, in a number of studies, heterogeneity may have been mislabelled as inefficiency.

The aim of this paper is to illustrate how relaxing the assumption of a unique technology may alter both the efficiency and elasticity estimates obtained. In order to illustrate this point we will use a translog production function and estimate a Latent Class Stochastic Frontier model (Kumbhakar and Orea, 2004). A two-class model is estimated and the results obtained differ substantially from the more conventional approaches both in terms of estimated elasticities and computed efficiency scores.

There are three main findings which emerge from this paper. First, the estimate of the full sample efficiency score increases by 27%-29% (17-18 p.p.) from 0.61-0.62 to 0.79. From a policy-perspective this suggests that heterogeneity may have been mislabelled as inefficiency. As a consequence, while inefficiencies remain sizeable and tackling inefficiencies may still be important, this, alone, is unlikely to lead to the dramatic changes in productivity required to change the African agricultural sector "beyond recognition".

Secondly, we also notice very different elasticities of inputs across classes, with this being particularly stark in the cases of oxen, fertilizer and labour. This is telling of the heterogeneity in farming in Ethiopia, with different technologies having very different potentials. In terms of policy, at the micro level, the results suggest that there is a need for differentiated policies which take into account the production technology.

Finally, upon inspecting the geographical class subdivision, we note that there is a stark subdivision at peasant-association level, with the vast majority of households in a given peasant association falling into one class. At the regional level, the distinction is not as clear but one class seems to be composed mainly of households from the Highlands (in Tigray and Amhara), whereas the second class mainly consists of households from Oromia and SSNP (Southern Nations, Nationalities and Peoples) regions. As such, this method provides a data-driven way to incorporate heterogeneity, without the need of atomizing studies to the village/peasant association level or assuming blunt technological, regional or ecological divisions.

The remainder of the paper is structured as follows. Sections 2 and 3 provide a succinct overview of the data, the approach and the methodologies. Section 4 presents the results. Finally, section 5 concludes.

2.2 Agriculture in Ethiopia, farm efficiency and the case for a Latent Class model

2.2.1 Agriculture in Ethiopia

The past and current importance of the agricultural sector in Ethiopia cannot be overstated, with some of the darkest moments of Ethiopia's recent history being closely associated with the performance of the agricultural sector, as epitomized by recurrent famines in the 1970s and 1980s, and more recently in 1999-2000 (Devereux, 2009, Headey et al., 2014). However, it is also the sector that holds the key to spurring economic development and reducing poverty, accounting for 73% of total national employment (World Bank WDI, 2016). Moreover, approximately 90% of the poor rely on agriculture as a source of livelihood (Yu et al., 2011). Cereals, in particular, are especially important crops, as highlighted by 2007 CSA (Central Statistics Authority) estimates, which placed 70% of all area under agricultural production under cereal production (Yu et al., 2011).

A major challenge will be to achieve increases in agricultural production and productivity in era characterized by fast population growth, often perceived as being responsible for land degradation (Taddese, 2001), and long-term declines in available per-capita agricultural area (Jayne et al., 2003), which may limit the potential expansion of the smallholder agricultural sector. The conjugation of both factors effectively implies not only that tackling the efficiency in the production of cereals is crucial, but also that it will become increasingly important in the foreseeable future.

2.2.2 Efficiency Analysis in Africa and Ethiopia: Motivating the use of a Latent Class Model

Thus, unsurprisingly, the analysis of efficiency in Ethiopia has attracted the attention of a number of authors. As will be explained below, the majority of the studies typically assume a unique technology for the whole sample. When this is not the case, authors tend to either divide the sample into blunt technological or regional divides.

However, in a country such as Ethiopia, which is characterized by widespread natural and technological heterogeneity, an estimation which fails to account for heterogeneity may be problematic and may lead to: 1) mislabelling heterogeneity as inefficiency, with households being compared to a potentially inappropriate aggregate frontier; and 2) a single set of input elasticity estimates, which may fail to portray the very diverse needs of farmers. As a consequence, policies based on analyses which disregard heterogeneity may, ironically, be an inefficient way to spur efficiency.

Intuitively, a stochastic frontier model assesses the efficiency, based on observable factors, by computing a ratio of the position (the observed output) of a decision-making unit (a household) relative to its computed production possibility frontier (the maximum achievable output for a given set of inputs). Simply put, a unique class implies that, conditional on the quantity of inputs used, every household is compared to the same frontier. As a result, if there are two (or more) different production "classes" with different input elasticities, they would still be compared to a unique production frontier. Their position relative to the frontier would be captured as inefficiency. In reality, however, the estimated inefficiency would be capturing a mix of inefficiency as well as differences (heterogeneity) in the actual potential of both groups. As a result, a certain household may be compared to an "inappropriate" frontier and, thus, the computed inefficiency could be hiding heterogeneity. At the aggregate level, the implications of this may have the effect of misrepresenting the feasible efficiency gains.

Secondly, as highlighted previously, a common production function may fail to represent the underlying diversity of production technologies. Taken to the extreme, it would be hard to argue against the fact that a household using a low quantity of inputs in drought-prone Tigray would require very different policies from a household engaging in higher-input agriculture in the more fertile areas of Oromia or Amhara. Not least because, agronomically, as pointed out by both Pender and Gebremedhin (2008) and Gebregziabher et al. (2012), fertilizer tends to perform more poorly in moisture-stressed environments in the absence of irrigation. However, in an attempt to obtain important information that is relevant for policy-making at the national level, it is not unusual in the technical efficiency literature in Sub-Saharan Africa to assume a common production function across wide geographical areas where the technologies may, or may not, be common. This type of analysis has been done in a number of countries including Malawi (Tchale, 2009) and Ethiopia (Bachewe et al., 2011, Mekonnen et al., 2013, Bachewe et al., 2015). Interestingly, these studies all tend to find relatively low efficiency levels, in the 0.37-0.6 range.

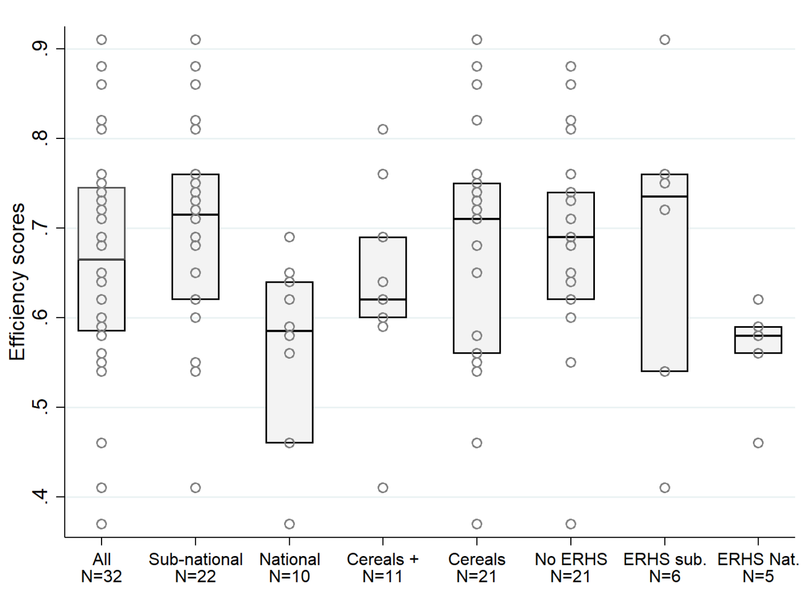
Figure 2.2 shows the distribution of average technical efficiency scores found in twenty-nine studies that estimated technical efficiency scores in Ethiopia and included cereals in their analysis[2], according to a number of characteristics of the study[3]. While the number of studies

[2]The list of studies is available in the Appendix of the thesis (Table 2A.1). I used the list of papers in Ogundari (2014) as a starting point. Specifically, I focused on the papers focusing on cereals and Ethiopia. Then, an additional search was made to expand the list of papers. A conscious effort was made to find papers that used the ERHS, as the calculated efficiency scores are more likely to be comparable to the results presented in this paper

[3]Figure 2.2 shows that 32 technical efficiency scores were used, despite there only being 29 studies. This occurs because certain studies use more than one different methodology to compute the efficiency scores. Consequently, in certain cases, more than one efficiency score was used per study.

analised is too small to claim that a pattern emerges, in general, studies using a large sample and performed at the national level tend to find lower efficiency scores than studies performed at the sub-national[4] level. Intuitively, an explanation for this could be that, in more confined geographical areas, it is more likely that the same (or a similar) production function actually holds for a given technology. This would also imply that it is more unlikely that heterogeneity is captured as inefficiency.

Figure 2.2: Boxplots of estimated technical efficiency scores in Ethiopia



Source: Author's calculations

Authors have recognized the importance of incorporating either the natural or the technological heterogeneity. The difficulty lies in deciding how, how much, and what type of heterogeneity to incorporate in the analysis. In order to fully capture heterogeneity in an accurate way, it would be tempting to atomize the study at the village or peasant association level. Doing so would imply that one could assume with relative safety that the assumption of a common production function holds (at least for a large portion of households). However, this may also not be desirable. Such studies would likely suffer from methodological issues such as sample size. And even when accurate measurements could be performed, each individual study would bear little policy relevance. Moreover, the frontier of a given peasant association may also not represent the actual maximum potential for households in this peasant

[4]Sub-national was defined as a study focusing on two or less regions of Ethiopia. However, in all but two cases, sub-national studies focus only on one region.

association. Perhaps there are indeed large inefficiencies and in such cases the frontier from households in comparable villages using similar technologies may be more appropriate.

However, the other extreme of amalgamating the majority of the producers under one technology, is likely to be policy-relevant in its geographical coverage, but perhaps less accurate and less representative of the diverse groups in the sample. The key is then to find a trade-off between incorporating heterogeneity in such a way that that the results are still of relevance for nationwide policy-making. The challenge is therefore to find groups of comparable households which plausibly may share a similar frontier. Previous authors have addressed different types of heterogeneity and, in general, they have used sample stratification to address this. Alemu et al., 2009 and Bachewe et al., 2011, for instance, divide the sample by agro-ecological zones in order to understand whether efficiency scores vary substantially across agro-ecological zones. Other authors have also incorporated blunt technological heterogeneity in the form of users *vs.* non-users of fertilizer (Bachewe et al., 2011) or irrigated *vs.* non-irrigated (Gebregziabher et al., 2012) plots. However, such sub-divisions are also not devoid of criticisms.

One criticism is that sharp technological divides are unlikely to be an accurate description of a technology as they define a technology uni-dimensionally. Different production technologies tend to be characterized by the intensity of a number of different inputs including, but not limited to, irrigation and fertilizer. Moreover, different farming technologies also tend to require and/or allow for different management strategies by farmers. As such, it is important to divide households into groups that share similar technologies in a way that recognizes that technologies are multidimensional.

A second potential criticism relates to a sharp division by regions. Splitting a sample by region has the advantage that, often, regional policies may be more similar within a region. However, an issue with this is that natural potential is not perfectly proxied by regional division. In our case, there are cases in which, although peasant associations are in the same region, they do not belong to the same agro-ecological zone. As such, it is also important to group households in a way that does not *a priori* impose sharp regional divides.

In order to capture the full picture, we would need to have sub-samples which are broadly comparable in terms of their technology. Therefore, latent classes may present a more suitable alternative to modelling heterogeneity. Latent class models are widely regarded as a more parsimonious way of representing heterogeneity than, for example, random-effects and fixed-effects models (Brown et al., 2014). In practice, as explained in Brown et al., 2014, latent classes split the population into a number of sub-samples, for which the same statistical model applies, but for which parameters may differ. Intuitively, latent class models allocate households who have different coefficients into different classes. This allows us to circumvent the issue of the sharp, uni-dimensional technological division as well as sharp geographical and ecological divides. This method therefore allows us to find a data-driven method to constitute classes that are broadly comparable. We can then account for the different potentials by agro-ecological zones by adding agro-ecological-year dummies.

However, it being a statistical method, the latent class model does not guarantee that the households pooled together in one class are grouped together in a way that makes sense. As such, after estimating the latent class model we will observe the spatial distribution of the classes so as to assess whether the subdivision that it implies makes sense. We would expect there to be a stark pattern at the peasant association level, where technologies are likely to be similar. As a result, for most cases, we would not expect to see households of the same village evenly split across classes. Instead, we would expect households in a given peasant association to be allocated majoritarily to the same class. In addition to this, since agro-ecological zones and regions capture, to some degree, natural potential and broad trends in agricultural policy, we would also expect to see a pattern emerging by agro-ecological zone and regions.

There are a number of papers which apply latent class models in order to divide samples into comparable groups. Alvarez and del Corral (2010) and Alvarez and Arias (2015) focus on cattle farming and divide the sample into intensive and extensive cattle farming. The authors use the average stocking rates and the average concentrate per cow as separating variables. Barath and Fertho (2015) also use a latent class model to crop producing farms in Hungary. Specifically, the authors use land size as the only explanatory variable determining class allocation. Finally, Sauer and Paul (2013) also use a latent class model and focus on the dairy sector in the EU. The authors focus on technical change and elasticities and use four separating variables, namely labour, fodder, milk as a fraction of output and organic output as a percentage of total revenue.

The approach used in this paper shares some similarities with these papers. For instance, in terms of the class allocation equation, the approach is quite similar to Sauer and Paul (2013), as it considers a number of different functions. In terms of the estimation procedure, the approach used in this paper is most similar to the approach used in Barath and Fertho (2015) and in Alvarez and del Corral (2010).

However, the approach chosen here also differs from the approaches used in other papers in a number of ways. First, by using a number of separating variables (such as in Sauer and Paul, 2013), this paper partly addresses the multidimensional aspect of agricultural technologies, which is missing in Barath and Fertho (2015). Second, other papers have typically used the Aigner model as a benchmark. However, they did not assess the latent class model against another benchmark model which modelled unobserved heterogeneity. Third, most of the literature using latent class models has focused on European agriculture, which is often seen as relatively efficient. For instance, Barath and Fertho (2015) find an average technical efficiency of 0.74 using the Aigner model, which is a higher efficiency score than the majority of studies focusing on African agriculture. As such, it is likely that the overestimate of potential gains is higher in the Ethiopian context.

In this paper we use five separating variables to account for different technologies. Specifically, we use the average land size and the average intensity of three inputs (Fertilizer, Labour and Draft power)[5] as well as the average proportion of area under cereals over the sample period. We believe that land size, together with the intensity with which inputs are used is a plausible way to capture the production technology used[6]. Proportion of area under cereal production is used in order to capture the degree of specialization or the importance given to cereals by a given household. Averages, rather than levels are used to keep class membership fixed over time (Brown et al., 2014).

Given the separating variables we use, there are a number of different possible class allocations. It could result in an intensive *vs.* extensive dichotomy or it could result in classes which use intensively different kind of inputs.

[5]Input intensity is constructed as the quantity of a given input divided by area under cereal production.
[6]Some systems are inherently more intensive in certain inputs.

2.3 Data, Methodology and Empirical Specification

2.3.1 Data

All the data used in this study comes directly from the Ethiopian Rural Household Survey[7] (ERHS, 2011) and all the waves from 1994[8]. Only farmers that cultivate over 0.01 hectares of cereals in every period were kept for the analysis. The 1989 wave was excluded as it is not comparable to subsequent waves in terms of its geographical coverage (Dercon and Hoddinott, 2004).

The 1994 wave is composed of 1,470 households from 18 different peasant associations (15 different villages[9]), spread over 4 regions. As mentioned by Dercon and Hoddinnot (2004), this sample is not nationally representative, or even fully representative of rural Ethiopia. The sample can be viewed as broadly representative of households in non-pastoralist farming systems as of 1994 (Dercon and Hoddinnot, 2004). Thus, when aggregated efficiency scores are presented by region, these refer solely to the units used in the sample and is not necessarily representative of Ethiopia. We use a slightly modified version of the agro-ecological division used by Bachewe et al. (2011), summarized in Table 2.1. The only modification is that we put together the Hararghe and the Arussi/Bale agro-ecological zones under "Other". As a result, we are left with four agro-ecological zones, namely the Northern Highlands, Central Highlands, Enset producing areas and Other.

[7] These 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.

[8] We use 6 waves, namely 1994, 1995, 1997, 1999, 2004 and 2009.

[9] As mentioned in Dercon and Hoddinott (2004), the largest village (Debre Berhan) is divided into four parts.

Table 2.1: List of Peasant Associations by AEZ

Agro-Ecological Zone	**Peasant Association**
Northern Highlands	Haresaw
	Geblen
	Shumsheha
Central Highlands	Dinki
	Debre Berhan Milki
	Debre Berhan Kormargefia
	Debre Berhan Karafino
	Debre Berhan Bokafia
	Yetmen
	Turufe Ketchema
Enset	Imdibir
	Aze-Deboa
	Adado
	Gara-Godo
	Do'oma
Other	Sirbana Godeti
	Korodegaga
	Adele Keke

Source: Adapted and changed from Bachewe et al., 2011

The panel is unbalanced but we will use only the balanced sample for the remainder of this paper. In order to highlight the differences in input-use across Ethiopia we provide summary statistics for the full sample as well as by agro-ecological zone in Table 2.2. As shown in the first two columns, the average farmer cultivates 1.18 hectares of cereals, allocate about 70%

of the total area to cereals and has a yield of approximately 838 kgs/ha. The mean level of fertilizer use on cereals is 52 kg.

However, there are stark differences across agro-ecological zones. As made clear by Table 2.2, yields are substantially higher in the parts of Amhara and Oromia which are part of the Central Highlands and Other agro-ecological zones. In addition, in these two agro-ecological zones, households tend to use larger quantities of fertilizer and cultivate larger plots. On the other hand, the Northern Highlands, which typically have lower amounts of rainfall have lower yields and use virtually no fertilizer. As mentioned by Gebrehiwot et al. (2011), most drought crises occurred in the Northern Highlands (which include all the peasant associations in Tigray and one peasant association in Amhara). Finally, in the Enset agro-ecological zone, smaller areas are devoted to cereals. Yields in this agro-ecological zone are higher than in the Northern Highlands but lower than in the Central Highlands and "Other" areas.

Table 2.2: Summary statistics (by agro-ecological zone)

	Full Sample		**N. Highlands**		**C. Highlands**		**Other**		**Enset**	
Variables	Mean	Std.Dev.	Mean	Std.Dev.	Mean	Std.Dev.	Mean	Std.Dev.	Mean	Std.Dev.
Cereal production (kg)	815.99	1001.73	364.31	442.05	1108.00	944.98	1381.11	1394.15	224.06	270.74
Cereal yield (kg/ha)	838.39	765.56	549.67	536.05	972.37	786.22	952.97	778.75	781.93	821.23
Cereal area (ha)	1.18	1.12	0.88	0.99	1.43	1.03	1.82	1.31	0.43	0.56
Cereal area (proportion)	0.70	0.26	0.86	0.20	0.71	0.22	0.78	0.19	0.46	0.28
Fertilizer used (kg)	52.88	82.31	3.51	12.20	80.87	87.43	88.16	109.41	19.88	28.58
Number of oxen	0.91	1.11	0.71	0.85	1.21	1.13	1.12	1.33	0.39	0.81
Household size	6.16	2.70	5.45	2.46	6.00	2.65	6.40	2.55	6.86	2.94
Tigray	0.12	0.33	0.59	0.49	0.00	0.00	0.00	0.00	0.00	0.00
Amhara	0.35	0.48	0.41	0.49	0.72	0.45	0.00	0.00	0.00	0.00
Oromya	0.31	0.46	0.00	0.00	0.28	0.45	1.00	0.00	0.00	0.00
SSNP	0.22	0.41	0.00	0.00	0.00	0.00	0.00	0.00	1.00	0.00
N	5034		1050		1830		1050		1104	
N Households	839		175		305		175		184	

N. Highlands refers to Northern Highlands. C. Highlands refers to central Highlands.

2.3.2 Methodology

Stochastic frontier model - Technical Efficiency, the Aigner et al. 1977 model and Fixed Effects

The measurement of technical efficiency defines a ratio between the observed value against the maximum potential (unobserved) value which can possibly be achieved by the unit of observation, given its level of inputs (Fried et al., 2008). Thus, measuring efficiency inevitably implies creating a frontier characterising the (unobserved) maximum quantity of output for a given level of inputs[10].

In the Stochastic Frontier Model approach, first proposed independently and simultaneously by Aigner et al. (1977) and Meeusen and van den Broeck (1977), the cannonical stochastic frontier model can be algebraically represented by equation 2.1:

$$y_{it} = exp\{f(x_{it};\beta)\} * exp\{\varepsilon_{it}\}; \quad \varepsilon_{it} = v_{it} - u_i \tag{2.1}$$

This essentially means that the output of a given household at time t is a function of a deterministic component $f(x_{it};\beta)$ (determined by the input coefficients) which, in our case, will be defined using a Cobb-Douglas functional form, and an error term ε_{it}. The error term ε_{it} is itself composed of a noise term v_{it} (normally distributed) - which accounts for factors such as shocks due to variations in the performance of inputs or weather —and an inefficiency term, u_i —which is non-negative (following a half-normal distribution) (Coelli et al., 2005). The degree of technical efficiency (TE) of a given unit can then be computed as the ratio given in equation 2.2.

$$TE = \frac{y_{it}}{exp\{f(x_{it};\beta + v_{it})\}} = \frac{exp\{f(x_{it};\beta) + v_{it} - u_i\}}{exp\{f(x_{it};\beta) + v_{it}\}} = exp\{-u_i\} \tag{2.2}$$

Equation 2.2 states that the level of technical efficiency of household i is determined by the ratio of its output to the output of a predicted output of a fully-efficient household using the same input vector (Coelli et al., 2005).

In terms of estimation, an additional important methodological point is raised by Battese (1997). In our sample, a large proportion of households have a zero quantity of two inputs (fertilizer and oxen). Some authors circumvent this issue by adding a small number or omitting the households with zero values altogether in order to obtain the elasticities. However, adding a small number may bias the elasticities of the inputs. As such, for fertilizer and oxen (which have a large number of zeros) we follow the method proposed by Battese (1997) to deal with zero input values. The method consists of creating a dummy variable indicating whether the household use the input or not and then replace zero-values by one before taking the logs. In

[10] Or, in the case of the input-oriented specification, a minimum level of inputs for a certain level of output.

this way we can obtain consistent estimates. In addition to this, in order to capture common changes in the dependent variable across time, we include year dummy variables. We opt for year dummy variables over time trends since they do not impose a specific shape in the relationship for the dependent variable over time. In addition, we interact the time dummy variables with the agro-ecological zones in order to allow changes over time to be common to a given agro-ecological zone, rather than the full sample.

Finally, in terms of the functional form, we opted for a translog specification. This specification will be used throughout the rest of the paper and it can be algebraically denoted by equation 2.3:

$$ln\, y_{it} = \alpha + \sum_{h=1}^{h=2} \beta_{dj} d^*_{jit} + \sum_{j=1}^{j=n} \beta_j ln\, x^*_{jit} + \frac{1}{2} \sum_{j=1}^{j=n} \sum_{k=1}^{k=n} \beta_{jk} ln\, x^*_{jit} * ln\, x^*_{kit} + \sum_{t=1}^{t=6} \sum_{a=1}^{a=4} \beta_{ta} d_t * d_a + v_{it} - u_{it} \tag{2.3}$$

In Equation 2.3 we model the natural logarithm of total production (y_{it}) as a function of a constant (α). We use a translog as it is more flexible and we do not need to impose *a priori* restrictions on the estimated technologies (Alvarez and Arias, 2015). One noticeable difference between the function we estimate and the typical translog function is our treatment of explanatory variables that have zero values (h). We follow the method proposed by Battese, 1997 and for those variable that have zero values (fertilizer and number of oxen), we add a dummy variable (equal to one when the quantity of the input is zero, and zero otherwise). For these variables x^* is defined as $x^* = max(d^*_{jit}, x_{jit})$. For the other variables (land size and household size) there is no transformation of the variable and no dummy variable (i.e.$x^*_{jit} = x_{jit}$). We also add a set of agro-ecological-zone dummies (d_a) interacted with time dummies (d_t), which allow non-parametric time-trends to vary across agro-ecological zones. The error term of the equation is composed of a statistical noise parameter v_{it} (i.i.d.) and inefficiency u_i, which follows a half-normal distribution. This model, also known as Pooled Stochastic Frontier model, will be one of the "baseline" models against which the Latent Class model will later be compared. Given that we are using a Translog specification, elasticities are estimated using the following formula:

$$E_{ij} = \beta_j + \beta_{jj} lnx_{ij} + \sum_{k \neq j}^{n} \beta_{jk} lnx_{ik} \tag{2.4}$$

This model will be one of our baseline specifications. In addition to this, we use a second baseline specification which includes fixed effects and allows for time-invariant heterogeneity across households (Fixed Effects) but not technological differences. The main idea behind the True Fixed Effects model is that a part of the inefficiency is, in fact, capturing time-invariant heterogeneity (Kumbhakar et al., 2015). A big weakness of this model, however, is that it is hard to distinguish between time-invariant heterogeneity and inefficiency. As such, this model is not fully consensual since a mix of inefficiency and heterogeneity may be captured and it is not always clear which prevails, as it often depends on context. In our case, we use this model to include a benchmark which incorporates heterogeneity in order to compare with the latent class estimates. Equation 2.5 denotes the fixed effects specification used.

$$ln\, y_{it} = \alpha_i + \sum_{h=1}^{h=2} \beta_{dj} d^*_{jit} + \sum_{j=1}^{j=n} \beta_j ln\, x^*_{jit} + \frac{1}{2} \sum_{j=1}^{j=n} \sum_{k=1}^{k=n} \beta_{jk} ln\, x^*_{jit} * ln\, x^*_{kit} + \sum_{t=1}^{t=6} \sum_{a=1}^{a=4} \beta_{ta} d_t * d_a + v_{it} - u_{it} \tag{2.5}$$

Where the addition of the term α_i, included in equation 2.5, represents time-invariant heterogeneity of household i.

Latent Class Model

In this class of models, the underlying assumption is that there are a number of groups for which the same statistical model applies, but who may have different coefficients (Brown et al. 2014). Following Orea and Kumbhakar (2004), we can represent the underlying model by equation 2.6 :

$$y_{it} = exp\{f(x_{it}; \beta_c)\} * exp\{v_{it}|_c\} * exp\{-u_{it}|_c\} \tag{2.6}$$

Where the main difference between equations 2.6 and 2.1 is the fact that all the parameters are conditional on the class c to which a given household belongs. This can be represented by equation 2.7:

$$lny_{it} = \alpha_c + \sum_{h=1}^{h=2} \beta_{cdj} d^*_{jit} + \sum_{j=1}^{j=n} \beta_{cj} lnx^*_{jit} + \frac{1}{2} \sum_{j=1}^{j=n} \sum_{k=1}^{k=n} \beta_{cjk} lnx^*_{jit} * lnx^*_{kit} + \sum_{t=1}^{t=6} \sum_{a=1}^{a=4} \beta_{cta} d_t * d_a + v_{it}|_c - u_{it}|_c \tag{2.7}$$

Equation 2.7 essentially states that the outcome of household i and time t also depends on the class, c, to which the household belongs. This is the case because the vector of estimated parameters for the time dummies and inputs are contingent on class, as are the noise and inefficiency terms. We opt for AEZ-year dummy variables since we do not want to assume a common evolution of the frontier over time for the full sample. Instead, we prefer to assume a common evolution within an agro-ecological zone. The benefit of doing this is that the agro-ecological zones are more likely to capture the potential of the technology in a given environmental setting. The disadvantage of doing this, however, is that while we are not imposing a strict ecological divide, we are imposing that there are at least enough households in each class from all the agro-ecological zones so that the agro-ecological-zone-year dummies can be estimated[11].

As mentioned in Alvarez and Arias (2015), when using a latent-class model we need to consider three likelihood functions. First, is the likelihood of a given household i at time t belonging to class c. This can be denoted as:

$$LF_{ict} = g(y_{it}, x_{it}, \theta_c) \tag{2.8}$$

Where θ_c represents the sets of parameters for class c and g represents the likelihood function of a production frontier.

Second, we need to consider the likelihood for houshold i conditional on class c. This can be obtained as the product of the likelihood functions for each period:

[11] In the appendix (Tables 2A.3-2A.5) we show the results obtained when only time dummies are included for each class.

$$LF_{ic} = \prod_{t=1}^{t=T} LF_{ict} \tag{2.9}$$

Finally, there is the unconditional likelihood of household i which weighs the conditional likelihood by class (equation 2.9) by the prior probabilities:

$$LF_i = \sum_{c=1}^{c=C} LF_{ic} * P_{ic} \tag{2.10}$$

where the prior probabilities (P_{ic}) can be interpreted as the probabilities attached to the membership in class c. These conditional probabilities which determine the class allocation are then parametrized by a multinomial logit model (Khumbakar and Orea, 2004) where the prior probability of individual i belonging to class c is given by:

$$P_{ic} = \frac{\exp(\gamma + \delta_c q_i)}{\sum_{c=1}^{c=C} exp(\gamma + \delta_c q_i)} \tag{2.11}$$

Where δ_c is a vector of estimated coefficients and q_i a set of time-invariant coefficients (in our case this refers to the average intensity of a given input over the sample period). The reason why generally time-invariant variables are preferred over time-varying ones is because class membership tends to be viewed as being fixed. Posterior probabilities can then be calculated as follows:

$$Pr_{ic} = \frac{LF_{ic} * P_{ic}}{\sum_{c=1}^{c=C} P_{ic} * LF_{ic}} \tag{2.12}$$

As previously stated, P_{ic} denotes the prior class probability, $P(i,t|c)$ denotes the probabilities of observation i conditional on class c in a given period, t. Following this, the most likely class c^* is estimated and the inefficiencies $u_i|_c$ are estimated. Mean efficiencies are estimated in the form of Jondrow et al. (1982) (Besstremyannaya, 2011).

2.4 Results

2.4.1 Full Sample

Each of the stochastic frontier models described in section 3 was then estimated and the results are presented in Table 2.3. The first column shows the results for the Aigner et al. (1977) model, the second column shows the results for the "True" fixed effects model (Greene, 2005) and the remaining columns show the results for the two-class Latent Class Model. In the case of the latent class model, the maximum number of classes that could be disentangled by the software was two (with the three-class model failing to converge in NLOGIT 5.0) .

As can be seen from the results from the first two columns of Table 2.3 (Pooled OLS and Fixed effects) the estimates obtained from using either model are very similar. As can be seen in Table 2.4, the Pooled OLS model and the Fixed Effects models both suggest elasticities for Cereal Area, Oxen, Labour and Fertilizer are of approximately 0.50, 0.20, 0.17 and 0.03, respectively, thus implying slightly decreasing returns to scale. In terms of the estimated full-sample elasticities, as shown in table 2.5, we find average efficiency scores in the 0.61-0.62 range. This number is consistent with previous studies (e.g. Mekonnen et al., 2013) focusing on Ethiopia, performed at the national level using the same dataset and using similar dependent variables. The Pooled OLS and Fixed effects results are similar in terms of the estimated coefficients and efficiency scores. However, the results emerging from the latent class model differ substantially.

Table 2.3: Results - Stochastic frontier model
Pooled OLS, Fixed effects and 2-class LCM

	Pooled OLS	F.E.	Latent Class	
			Class 1	Class 2
Constant	5.810***		6.685***	5.262***
	(0.116)		(0.149)	(0.185)
Area	0.476***	0.500***	0.295***	0.729***
	(0.040)	(0.026)	(0.058)	(0.056)
Oxen	0.553***	0.517***	0.069	0.780***
	(0.132)	(0.095)	(0.170)	(0.182)
Household size	-0.102	-0.174***	-0.066	0.039
	(0.082)	(0.049)	(0.099)	(0.116)
Fertilizer	-0.306***	-0.393***	-0.200***	-0.261***
	(0.042)	(0.262)	(0.051)	(0.067)
Area sq	-0.074***	-0.069***	-0.113***	-0.017
	(0.014)	(0.009)	(0.021)	(0.020)
Oxen sq	0.115	0.140	0.293*	-0.158
	(0.130)	(0.093)	(0.166)	(0.173)
Household size sq	0.158***	-0.195***	0.209***	-0.013
	(0.053)	(0.033)	(0.059)	(0.078)
Fertilizer sq	0.141***	0.169***	0.099***	0.099***
	(0.012)	(0.008)	(0.014)	(0.020)
Area * oxen	0.052	0.0485*	-0.027	0.126***
	(0.037)	(0.027)	(0.055)	(0.048)
Area * Household size	0.021	0.019	0.092***	-0.046
	(0.021)	(0.014)	(0.026)	(0.030)
Area * Fertilizer	-0.022***	-0.030***	-0.001	-0.024**
	(0.006)	(0.004)	(0.008)	(0.011)
Oxen * Household size	-0.084	-0.068	0.055	-0.189**
	(0.065)	(0.046)	(0.074)	(0.096)
Oxen * Fertilizer	-0.092***	-0.095***	-0.068***	-0.048**
	(0.014)	(0.011)	(0.019)	(0.020)
Household size * Fertilizer	0.011	0.014*	-0.023*	0.048***
	(0.011)	(0.007)	(0.013)	(0.017)
Dummy fertilizer	-0.377***	-0.483***	-0.389***	-0.232*
	(0.068)	(0.041)	(0.076)	(0.119)
Dummy oxen	-0.014	-0.017	-0.092***	-0.064
	(0.027)	(0.019)	(0.031)	(0.039)
Year-AEZ FE	Yes	Yes	Yes	Yes
			Selection equation	
Constant			1.605**	
			(0.782)	
Av. Proportion cereal			-2.549***	
			(0.800)	
Av. Area			-0.490***	
			(0.188)	
Av. Labour intensity			-0.029**	
			(0.012)	
Av. Fertilizer intensity			0.022***	
			(0.003)	
Av. Oxen intensity			-0.403***	
			(0.117)	
Prior Probabilities			0.539	0.461
		Variance parameters		
Lambda	1.171***	1.013**	2.003***	0.802***
	(0.042)	(0.033)	(0.185)	(0.216)
Sigma	0.859***	0.888***	0.781***	0.716***
	(0.000)	(0.009)	(0.023)	(0.041)
AIC/N	2.086	2.175	1.888	
Log Likelihood	-5207.616	-4593.615	-4662.897	

All the variables are in natural logarithms with the exception of the dummy variables. Number in parentheses denote standard errors. *, **, *** denote statistical significance at the 10%, 5% and 1% level, respectively.

Table 2.4: Input elasticities Translog

Input	OLS		FE		Class 1		Class 2	
Variables	Mean	S. D.	Mean	S. D.	Mean	S. D.	Mean	S. D.
Land	0.50	0.10	0.50	0.11	0.50	0.13	0.64	0.06
Oxen	0.20	0.19	0.19	0.19	0.04	0.15	0.31	0.17
Labour	0.17	0.09	0.17	0.11	0.20	0.15	0.07	0.10
Fertilizer	0.03	0.29	0.02	0.34	0.04	0.19	-0.01	0.20

S. D. refers to Standard deviations

Table 2.5: Summary statistics - Efficiency scores by agro-ecological zone

All				
	Mean	**S.D.**	**Min.**	**Max.**
Pooled OLS	0.61	0.13	0.15	0.90
Fixed Effects	0.62	0.09	0.21	0.87
LCM (2 classes)	0.79	0.10	0.20	0.96
Northern Highlands (N=175)				
	Mean	**S.D.**	**Min.**	**Max.**
Pooled OLS	0.60	0.13	0.23	0.88
Fixed Effects	0.62	0.10	0.27	0.83
LCM (2 classes)	0.81	0.08	0.39	0.94
Central Highlands (N=305)				
	Mean	**S.D.**	**Min.**	**Max.**
Pooled OLS	0.61	0.12	0.15	0.86
Fixed Effects	0.63	0.08	0.23	0.81
LCM (2 classes)	0.80	0.10	0.20	0.94
Others (N=175)				
	Mean	**S.D.**	**Min.**	**Max.**
Pooled OLS	0.61	0.12	0.18	0.90
Fixed Effects	0.62	0.09	0.26	0.87
LCM (2 classes)	0.79	0.10	0.28	0.96
Enset (N=184)				
	Mean	**S.D.**	**Min.**	**Max.**
Pooled OLS	0.61	0.13	0.15	0.87
Fixed Effects	0.62	0.10	0.21	0.84
LCM (2 classes)	0.77	0.12	0.29	0.96

S. D. refers to standard deviations. Min. refers to the minimum value. Max. refers to the maximum value

Table 2.6 provides summary statistics for a number of variables by class membership. One noticeable factor is that input intensities differ significantly across classes. For instance, households in class one tend to use fertilizer a lot more intensively, cultivate less land (1 ha, on average) and use labour slightly more intensively. The intensity of use of oxen, however, is very similar in both classes. In this paper we thus interpret households in class one as the households using more intensive agriculture (especially in fertilizer) and those in class two as households engaged in more extensive (or lower-input) agriculture[12].

Table 2.6: Summary statistics by latent class (2 class translog)

	Full sample		Class 1		Class 2	
Variables	**Mean**	**S. D.**	**Mean**	**S. D.**	**Mean**	**S. D.**
Cereal production (kg)	815.99	1001.73	940.31	970.88	701.62	1016.12
Cereal yield (kg/ha)	838.39	765.56	1074.61	848.42	621.09	603.91
Cereal area (ha)	1.18	1.12	1.00	0.91	1.34	1.27
Cereal area (proportion)	0.70	0.26	0.63	0.27	0.77	0.25
Fertilizer used (kg)	52.88	82.31	69.74	95.08	37.37	64.73
Number of oxen	0.91	1.11	0.83	1.09	0.98	1.12
Household size	6.16	2.70	6.40	2.86	5.93	2.53
Labour intensity (people/ha)	17.01	42.54	18.27	30.01	15.86	51.41
Fertilizer intensity (kg/ha)	57.72	112.83	91.46	148.68	26.69	45.85
Oxen intensity (oxen/ha)	1.20	2.88	1.19	2.46	1.22	3.23
Tigray	0.12	0.33	0.00	0.05	0.24	0.42
Amhara	0.35	0.48	0.27	0.44	0.42	0.49
Oromya	0.31	0.46	0.43	0.50	0.20	0.40
SSNP	0.22	0.41	0.30	0.46	0.14	0.35
Northern Highlands	0.21	0.41	0.12	0.33	0.29	0.45
Central Highlands	0.36	0.48	0.35	0.48	0.38	0.48
Other	0.21	0.41	0.23	0.42	0.19	0.39
Enset	0.22	0.41	0.30	0.46	0.14	0.35
N	839		402		437	
Proportion	100.00		47.91		52.09	

S. D. refers to Standard deviations. N refers to the number of households.

[12] Class two incorporates both extensive farms as well as low input households in the Northern Highlands. As such it includes both extensive agriculture and low-input agriculture.

The elasticities also appear to differ substantially across classes with households in the intensive class (class one) displaying higher elasticities of fertilizer and labour, whereas the extensive households (class two) display higher elasticities of land and oxen. With respect to the estimated efficiency scores, the two-class stochastic frontier model also highlights a very different pattern. As shown in table 2.5, while the efficiency scores estimated using our baseline specifications are in the 0.61-0.62 range, the average efficiency score estimated using the two-class LCM is 0.79. As such, while the first two methods would suggest that using inputs efficiently could potentially lead to increases in production approximately in the 61%-64% range, the two-class LCM estimates a much more modest potential increase of about 27%. As highlighted by table 2.5, this pattern is quite similar across agro-ecological zones with the baseline estimates all in the 0.60-0.63 range and the two-class LCM efficiency estimates in the 0.77-0.81 range.

In order to assess whether the class subdivision is plausible, apart from the differences in summary statistics, we provide a summary of the geographical location of households according to their class. Table 2.7 summarizes the number and proportion of households from a given village (peasant association) in a given class. The most salient feature of the table is that in all but one case (Shumsheha) 80% or more of the households in a village are allocated to the same class[13]. Intuitively, we would expect the majority of households in a given village to use similar technologies. As such, a stark subdivision at the village level is to be expected. Moreover, as shown by Table 2.6, there is also a relatively stark pattern emerging at the regional level, with the first class being predominantly (73%) composed of households from Oromia and the SSNP regions. In contrast, class two is predominantly (66%) composed of households of households from Tigray and Amhara. The subdivision by agro-ecological zone, however, is a lot less stark. This can be explained by the inclusion of year-agro-ecological zone dummy variables which forces households from each agro-ecological zone to be included in each class[14]. However, we believe that assuming a common trend for all households in a class without taking into account agro-ecological setting is not defensible, especially when cropping patterns and environmental conditions differ so widely by region.

[13]The four peasant associations in Debre Berhan are in the same village (Hoddinot and Dercon, 2004). In 10 out of 15 cases this percentage exceeds 90%.

[14]In the Appendix (Tables 2A.2 -2A.5) we present the results of a model which does not include AEZ-year dummy variables and the division by agro-ecological zone is a lot starker. In terms of the estimated efficiency scores do not change substantially, with average efficiency scores increasing from 0.55-0.57 (baseline) to 0.76 (LCM) for the full sample (Table 2A.4)

Table 2.7: Class allocation by peasant association (2 class lcm translog)

	Full sample	Class 1		Class 2	
Peasant Association	**N**	**N**	**Proportion**	**N**	**Proportion**
Haresaw	51	0	0.00	51	100.00
Geblen	53	1	1.89	52	98.11
Dinki	46	1	2.17	45	97.83
Yetmen	36	32	88.89	4	11.11
Shumsheha	71	49	69.01	22	30.99
Sirbana Godeti	60	56	93.33	4	6.67
Adele Keke	36	34	94.44	2	5.56
Korodegaga	79	1	1.27	78	98.73
Trirufe Ketchema	84	82	97.62	2	2.38
Aze Deboa	51	47	92.16	4	7.84
Adado	28	0	0.00	28	100.00
Gara Godo	73	69	94.52	4	5.48
Doma	32	5	15.63	27	84.37
Debre Berhan Milki	45	8	17.78	37	82.22
Debre Berhan Kormargefia	47	7	14.89	40	85.11
Debre Berhan Karafino	27	5	18.52	22	81.48
Debre Berhan Bokafia	20	5	25.00	15	75.00
Total	839	402	47.91	437	52.09

N refers to the total number of households. The numbers in the proportion column were computed by dividing the number of households from a given peasant association in a given class by the total number of households from that peasant association in our sample. All proportions are rounded to two decimal places.

2.5 Conclusion

Currently, there is somewhat of an incoherence in the policy discourse which often tends to discourage a "one-size-fits-all" approach and the way in which the Stochastic Frontier tends to be applied. While the policy discourse recognizes substantial heterogeneity across African smallholders, the way in which stochastic frontiers are often estimated tends to assume a common technology among all the units in the sample or sub-sample of interest.

This paper, through disaggregating technology by latent classes and computing separate production functions, provides evidence suggesting that a large proportion of the "inefficiency" found in the efficiency literature focusing on African countries may be partially capturing a large degree of heterogeneity in both the natural endowments, production technologies and their respective interactions. It then showed that slightly relaxing the common production technology assumption by allowing two groups, tended to produce groups of households which differed substantially in their input choice. In addition to this, the computed efficiency scores also increased substantially.

In particular, unweighted average efficiency scores in Ethiopia increase from around 0.61-0.62 to 0.79 in our preferred set of results[15]. In practice, it implies that the potential gains from tackling inefficiencies are reduced from about 61%-64% in the baseline scenario to approximately 27% in the two-class latent class specification. In addition to this, the analysis also highlighted the fact that elasticity estimates differ considerably across different classes.

The implications of this are far-reaching. First, from a policy perspective, the efficiency gains achievable from tackling the inefficiencies in input use may have been substantially overstated. As a consequence, this may imply that simply tackling inefficiencies in production is unlikely to change African Agriculture "beyond recognition". Second, it suggests that technological heterogeneity should be mirrored in agricultural policy, since elasticities differ substantially depending on the technology used. Households engaged in more intensive agriculture typically have larger elasticities of fertilizer and labour whereas households engaged in more extensive (lower-input) agriculture tend to have higher elasticities of land and oxen.

Moreover, this is unlikely to be an issue confined to Ethiopia and, hence, this could have far-reaching implications for agricultural policy in Africa. Therefore this provides two potential avenues for future research. The first would be to investigate whether the pattern found in this paper seems to be found in other African countries. Secondly, given the narrow focus of this paper, one limitation is that it ignores a number of important institutional factors such as human capital and institutions. It would be interesting to understand whether these affect class allocation and whether they have different effects on production, depending on the technology used.

[15]Increases from 0.55-0.57 to 0.76 in our alternative specifications in the appendix

References

[1] Aigner, D. L., Lovell, C. K., & Schmidt, P. (1977). Formulation and estimation of stochastic production function models *Journal of Econometrics*, 6(1), 21-37.

[2] Alemu, B. A., Nuppenau, E. A., & Boland, H. (2009). Technical efficiency of farming systems across agro-ecological zones in Ethiopia: An application of stochastic frontier analysis. *Agricultural Journal, 4(4).*

[3] Alene, A. D., & Hassan, R. M. (2003). The determinants of farm-level technical efficiency among adopters of improved maize production technology in western Ethiopia. *Agrekon*, 42(1), 1-14.

[4] Alvarez, A., & Arias, C., (2015). Effects of switching between production systems in dairy farming. *Bio-Based And Applied Economics*, 4(1), 1-16. doi:10.13128/BAE-14715

[5] Alvarez, A., & del Corral, J. (2010). Identifying different technologies using a latent class model: extensive versus intensive dairy farms). *European Review of Agricultural Economics*, 37(2), 231-250.

[6] Bachewe, F. N., Koru, B., & Taffesse, A. S. (2015). Cereal productivity and its drivers: The case of Ethiopia. *ESSP II Working Paper 75.* Addis Ababa, Ethiopia and Washington, D.C.: Ethiopian Development Research Institute (EDRI) and International Food Policy Research Institute (IFPRI). http://ebrary.ifpri.org/cdm/ref/collection/p15738coll2/id/129230

[7] Bachewe, F., Guush, B., Sinafikeh, A., Gerawork, G., Taffesse, A. S., & Hoddinott, J. (2011). Sources of Inefficiency and Growth in Agricultural Output in subsistence Agriculture: A stochastic Frontier Analysis.*International Food Policy and Research Institute,*

ESSP II Working Paper, (19).

[8] Barath, L., & Ferto, I. (2015). Heterogeneous technology, scale of land use and technical efficiency: The case of Hungarian crop farms. *Land Use Policy*, 42, 141-150.

[9] Battese, G. E. (1997). A note on the estimation of CobbDouglas production functions when some explanatory variables have zero values. *Journal of agricultural Economics*, 48(13), 250-252.

[10] Besstremyannaya, G. (2011). Managerial performance and cost efficiency of Japanese local public hospitals: A latent class stochastic frontier model. *Health economics*, 20(S1), 19-34.

[11] Brown, S., Greene, W. H., & Harris, M. N. (2014). A new formulation for latent class models.

[12] Coelli, T. J., Rao, D. S. P., O'Donnell, C. J., & Battese, G. E. (2005). An introduction to efficiency and productivity analysis. Springer Science and Business Media.

[13] Collier, P., & Dercon, S. (2014). African Agriculture in 50Years: Smallholders in a Rapidly Changing World?. *World Development*, 63, 92-101.

[14] Croppenstedt, A., & Muller, C. (2000). The impact of farmers' health and nutritional status on their productivity and efficiency: evidence from Ethiopia. *Economic Development and Cultural Change*, 48(3), 475-502.

[15] Devereux, S. (2009). Why does famine persist in Africa?. *Food security*, 1(1), 25-35.

[16] Dercon, S., & Hoddinott, J. (2004). The Ethiopian rural household surveys: Introduction. International Food Policy Research Institute, Washington, DC Photocopy.

[17] Ethiopia Rural Household Survey Dataset, 1989-2009. (2011). Washington, D.C. International Food Policy and Research Institute (IFPRI) (datasets).

[18] Fried, H. O., Lovell, K. C. A., & Schmidt, S. S. (2008). The Measurement of Productive Efficiency and Productivity Growth, USA: OUP

[19] Gebregziabher, G., Namara, R. E., & Holden, S. (2012). Technical efficiency of irrigated and rain-fed smallholder agriculture in Tigray, Ethiopia: A comparative stochastic frontier production function analysis. *Quarterly Journal of International Agriculture*, 51(3), 203.

[20] Gebrehiwot, T., van der Veen, A., & Maathuis, B. (2011). Spatial and temporal assessment of drought in the Northern highlands of Ethiopia. *International Journal of Applied Earth Observation and Geoinformation*, 13(3), 309-321.

[21] Greene, W. (2005). Reconsidering heterogeneity in panel data estimators of the stochastic frontier model. *Journal of Econometrics*, 126, 269-303.

[22] Headey, D., Dereje, M., & Taffesse, A. S. (2014) Land constraints and agricultural intensification in Ethiopia: A village-level analysis of high-potential areas. *Food Policy*, 48, 129-141.

[23] Jayne, T. S., Yamano, T., Weber, M. T., Tschirley, D., Benfica, R., Chapoto, A., & Zulu, B. (2003). Smallholder income and land distribution in Africa: implications for poverty reduction strategies. *Food policy*, 28(3), 253-275.

[24] Jondrow, J., Lovell, C. K., Materov, I. S., & Schmidt, P. (1982). On the estimation of technical inefficiency in the stochastic frontier production function model. *Journal of econometrics*, 19(2-3), 233-238.

[25] Kumbhakar, S. C., & Orea, L. (2004). Efficiency measurement using a latent class stochastic frontier model. *Empirical Economics*, 29(1), 169-183.

[26] Kumbhakar, S. C., Wang, H. J., & Horncastle, A. (2015). A Practitioner's Guide to Stochastic Frontier Analysis Using Stata. Cambridge University Press.

[27] Meeusen, W., & Van den Broeck, J. (1977). Efficiency estimation from Cobb-Douglas production functions with composed error. *International economic review*, 435-444.

[28] Mekonnen, D. K., Dorfman, J. H., & Fonsah, E. G. (2013). Productivity and Efficiency of Small Scale Agriculture in Ethiopia. In 2013 Annual Meeting, February 2-5, 2013, Orlando, Florida (No. 143038). Southern Agricultural Economics Association.

[29] Ogundari, K. (2014). The Paradigm of Agricultural Efficiency and its implications on Food Security in Africa: What Does Meta-analysis Reveal? *World Development*, 64, 690-702.

[30] Pender, J., & Gebremedhin, B. (2008). Determinants of agricultural and land management practices and impacts on crop production and household income in the highlands of

Tigray, Ethiopia. *Journal of African Economies*, 17(3), 395-450.

[31] Sauer, J., & Paul, C. J. M. (2013). The empirical identification of heterogeneous technologies and technical change. *Applied Economics*, 45(11), 1461-1479.

[32] Seyoum, E. T., Battese, G. E., & Fleming, E. M. (1998). Technical efficiency and productivity of maize producers in eastern Ethiopia: a study of farmers within and outside the Sasakawa-Global 2000 project. *Agricultural economics*, *19(3)*, 341-348.

[33] Taddese, G. (2001). Land degradation: a challenge to Ethiopia. *Environmental management*, 27(6), 815-824.

[34] Tchale, H. (2009). The efficiency of smallholder agriculture in Malawi. *African Journal of Agriculture and Resource Economics*, 3(2), 101-121.

[35] Tesfay, G., Ruben, R., Pender, J., & Kuyvenhoven, A. (2005). Resource use efficiency on own and sharecropped plots in Northern Ethiopia: determinants and implications for sustainability.

[36] World Bank (2016). World Development Indicators.

[37] Yu, B., Nin-Pratt, A., Funes, J., & Gemessa, S. A. (2011). Cereal production and technology adoption in Ethiopia. *Ethiopia Strategy Support Program II (ESSP II), ESSP II Working Paper*, 31.

Chapter 3

Revisiting the link between cereal diversity and production in Ethiopia

Abstract

Recently, a number of studies suggest that, at the micro level, cereal diversity positively affects mean yields and decreases the variability of yield. However, most analyses: 1) focus on very specific sub-regions, 2) use cross-sectional data; and 3) do not focus on the mechanism driving the effect. In this paper we revisit the link between cereal diversity and production using data from the Ethiopian Rural Household Survey. To this end, we use a mix of parametric and semi-parametric regression techniques.

For the full sample, we find a positive and significant effect of greater cereal diversity on cereal production. However, this positive effect seems to be driven by specific agro-ecological zones and by households who cultivate one crop in particular (teff), with these results being consistent using both parametric and semi-parametric methods. Overall, these findings indicate that 1) the scope for cereal diversity to drive increases in output may be limited, and 2) differences in potential yields from cereals in the crop mix seem to be part of the explanation. As a result, alternative conservation solutions may be needed.

JEL classification: Q16, Q57

3.1 Introduction

The effect of crop diversity on agricultural production remains somewhat of a conundrum in the Agricultural Economics literature, with recent micro-evidence at odds with historical trends in agricultural development. However, the importance of understanding this link remains important, as it speaks directly to the productivity of agricultural systems and their resilience to climate and weather. This implies that crop diversity may have an important role to play in terms of food security. However, whether increases in crop diversity represent a viable development strategy capable of delivering sustained productivity gains remains an open question.

Recently, a number of microeconomic studies seem to suggest a "win-win-win" situation in the form of increased productivity, reduced volatility of output and greater *in situ* conservation (Di Falco and Perrings 2003, Di Falco and Chavas 2006 and Di Falco et al. 2007). These findings, however, contrast sharply with historical trends in agricultural development, which appear to be driven by increasingly mechanized, specialized and input-intensive agriculture. This trend has been epitomized, at different periods in history, by cases such as the United States, Europe, and more recently, by China and India (Borlaug 2000, Evenson and Gollin 2003).

As such, a "micro-macro" paradox seems to have emerged. Studies at the farmer level seem to suggest that crop diversity positively affects agricultural productivity. At the macro level, however, increases in productivity in the most recent success stories seem to have been driven by less diverse systems. This current state of affairs is likely to be puzzling for policy-makers and is particularly important in the African context. According to Collier and Dercon (2014), the current African experience is unlikely to lead to the radical transformation of the agricultural sector, which could spur broad-based economic development. This implies that alternatives to the current model have to be sought.

Consequently, this paper looks at cereal diversity and seeks to understand whether an increase in cereal diversity represents a viable alternative leading to sustained productivity gains. Specifically, we focus on two questions. First, we test whether increases in crop diversity lead to productivity increases. Second, whether these effects can be explained by regional patterns and/or by a specific crop, indicating a composition effect.

In order to address these questions, we use data from the Ethiopian Rural Household Survey (ERHS) and use a mix of parametric and semi-parametric methods. Addressing this question in the context of Ethiopia using panel data is relevant since 1) agriculture has been selected to be an engine of socio-economic transformation (World Bank, 2007a); 2) much of the previous literature on crop diversity has focused on Ethiopia (Di Falco and Chavas, 2009; Di Falco et al., 2007; Di Falco et al., 2010; and Di Falco and Chavas, 2012a); and 3) the use of panel data is likely to mitigate concerns surrounding results from previous work.

Overall, two main findings emerge from this paper. First, consistent with previous literature, we find a sizeable average positive effect of crop diversity using both parametric and non-parametric methods. However, we also find this positive effect is found in a restricted set of agro-ecological zones. Second, we test whether this relationship could be driven by differences in the potential yield of the cereals in the crop mix. One of the cereals, teff, is notoriously a low-productivity crop (Vandercasteelen et al., 2013) and it could be driving the crop diversity results. Our results show that, when teff producers are excluded, the effects of crop diversity on production become noticeably smaller and insignificant in all parametric results. A similar conclusion is drawn from the semi-parametric results.

Overall, these results suggest that the positive diversity-productivity link could be weaker than suggested previously. Consequently, the scope for production gains from higher levels of crop diversity may be lower than previously thought. Furthermore, our results also suggest that a compositional effect, rather than the traditional complementarity and facilitation effects found in the ecological literature, could partly explain the positive relationship found in previous studies. Taken together, this questions the potential of increasing cereal diversity as a means to increase cereal productivity.

The rest of this paper is structured as follows. The next section provides a brief overview of the literature on farm biodiversity. Section three discusses the channels through which crop diversity may impact agricultural productivity. Section four discusses our measurement of crop diversity. Section five describes the data and the methodology used. Section six presents the results and section seven concludes.

3.2 Agriculture in Africa and Ethiopia

The current African experience promoting smallholder agriculture has not yet led to the productivity increases that will change African agriculture beyond recognition (Collier and Dercon, 2014). According to the authors, a radical transformation of the agricultural sector is deemed crucial for successful economic development. However, this transformation will have to occur in a very challenging environment defined by rapid demographic pressures as well as the looming threat of climate change. According to the UN world population prospects 2015, over half of the global population increase will occur in Africa. This, allied to potentially large losses arising from climate change (Schlenker and Lobbel, 2010), will make for a very challenging setting in which increases in productivity will need to happen.

In Ethiopia, the importance of the agricultural sector for its economic development is well recognized. As explained by Dercon and Zetlin (2009), since the early 1990s, the Ethiopian Governments growth strategy made the agricultural sector a pillar of its national development strategy, under the agricultural development-led industrialization (ADLI). This policy focused primarily on smallholders and, according to Rahmato (2008), sought to increase crop production through the provision and distribution of a number of modern inputs (including seeds and fertilizer) and training.

As a result, our sample period (1994-2009) was characterized by a rapid expansion in cereal area cultivated (World Bank 2015) and a strong growth in terms of agricultural output. However, the growth in cereal yields was more modest (World Bank 2007a and World Bank 2007b) and this was partly attributable to both land degradation and weather variability, which were found to have non-negligible effects. Since 2008, however, national-level data shows a significant increase in cereal yields from 1.45 tonnes/ha in 2008 to 2.33 tonnes/ha in 2014.

Currently, the agricultural sector remains vital. In 2013, Agriculture still accounted for about three quarters of total employment (73%) and 41% of GDP (World Bank, 2015). Looking forward, one key debate relates to whether production systems should be geared towards the traditional pattern of input-intensification or whether systems that are more diverse should be promoted (higher agro-intensification). This debate hinges directly on the link between increased crop diversity and production.

3.3 Crop diversity and productivity

In recent years there has been an increase in the number of studies that have looked at the link between various forms of biodiversity, including cereal diversity, and agricultural outcomes. In general, crop diversity may be beneficial for agriculture and development for a number of reasons.

From an ecological perspective, higher levels of on-farm crop diversity potentially represent an effective way of preserving plant genetic resources (Di Falco 2012). However, there are also a number of channels through which it may directly affect crop production directly.

The first such channel is through a sampling effect. In essence, a sampling effect implies that the higher the species richness (i.e. higher number of species), the larger the probability that the key species with the highest effects on the performance are present in the ecosystem (Tilman et al., 2005; Di Falco, 2012).

A second channel, as explained in Hooper et al. (2005) relates to a potential complementarity between crops. Different species use different resources at different times. Therefore, combining species which have different resource patterns may allow for such a complementarity effect, which is likely to result in a more efficient use of resources over time. As a result, in cases where resources are a limiting factor to growth and productivity, increasing the richness of the ecosystem could lead to greater productivity.

A third channel relates to a facilitation effect. This effect refers to positive interactions between species. An example of this effect can be found if, for example, one species is capable of providing a critical resource for other species or alleviate harsh environmental conditions (Hooper et al., 2005; Di Falco, 2012). According to Hooper et al. (2005) the complementarity and facilitation effects are two of the main reasons leading to overyielding (i.e. yields from a mixture of crops exceeding those of monoculture).

From an economic perspective, there are also a number of reasons why higher levels of agro-biodiversity may be desirable. As argued by Baumgärtner (2007), biodiversity has the potential to act as a natural insurance for risk-averse farmers, thus potentially being a substitute for financial insurance (Baumgärtner, 2007; Baumgärtner and Quaas, 2008; and Quaas and Baumgärtner, 2008). Moreover, as argued by Di Falco (2012), it allows farmers to produce and market their crops multiple times a year. This has the potential to hedge farmers against crop price volatility, as well as provide a smoother inflow of income.

Empirically, the majority of the evidence supporting increased crop diversity as a key source of productivity comes from studies in ecology performed in an experimental setting. Consequently, the experimental results need not translate to non-experimental settings where conditions are likely to differ substantially. This has led to the productive importance of biodiversity in agriculture being increasingly studied in non-experimental settings. The overarching results of this literature, however, seem to broadly corroborate the overall findings from the agroecology literature. The vast majority of studies focusing at the household level tend to find non-negligible economic gains from more diverse systems, both in the form of increased mean yields and reduced output volatility.

Evidence from Asia (Smale et al., 1998; Smale, 2008) as well as Europe (Di Falco and Chavas, 2006; Di Falco and Perrings, 2003; Di Falco and Perrings, 2005) all seem to suggest that higher levels of crop diversity are generally correlated with higher yields and lower variance in yields and/or reduced probability of crop failure. An additional study by Omer et al. (2007), which uses a stochastic frontier model approach, finds that higher levels of biodiversity are associated with a higher frontier and reduced inefficiency in the case of the UK.

In Africa, Ethiopia has probably been the most studied country and most of the research has focused on the Highlands of Ethiopia. In Tigray, Di Falco and Chavas (2009), Di Falco et al. (2007) and Di Falco and Chavas (2012a) all find that, on average, higher levels of crop diversity have a net positive effect on productivity. However, Di Falco and Chavas (2012a) highlight that there may be different sources of value for diversity. In particular, they find a positive complementarity effect (positive synergies between crops) and a negative convexity effect (scale effect). The latter provides an incentive to specialize. Overall, the authors still find a positive value of crop diversity. In the Amhara region, Di Falco et al. (2010), Di Falco and Chavas (2012b) and BangwayoSkeete et al. (2012) all find a positive effect of crop diversity on mean yield. In addition to this, Di Falco et al. (2010) also finds that this effect tends to be stronger when rainfall is lower.

In sum, most studies seem to suggest a positive effect of greater crop diversity on production, productivity and reduced variability. Moreover, the estimated effects also tend to be large, with Di Falco and Chavas (2012b), for example, finding an estimated effect of crop diversification amounting to approximately 17% of revenue for an average farm.

However, despite recent empirical evidence, a number of gaps remain in this literature. First, the majority of the literature focusing on Ethiopia has focused on the Ethiopian Highlands. As a result, findings may not be transposable to other areas of Ethiopia. Since the Ethiopian Highlands tend to be quite moisture-strained, it may be the case that this reduces the effectiveness of other inputs[1], thus favouring increasing crop diversity as an alternative. As a result, whether crop diversity yields similar gains across agro-ecological settings is still an open question. Secondly, previous research studying the cereal diversity-productivity link in Agricultural Economics does not convincingly answer why a positive relationship exists. Beyond the marginal effects, few studies have tried to disentangle the underlying mechanism driving this link. This is not a trivial question since policy implications will differ depending on whether the result is driven by one specific crop (a "sampling" effect) or whether it is driven by interactions between cereals ("complementarity" and "facilitation effects"). Previous typically does not test for the possibility that results could be driven by the inclusion of particular high- or low- performing crop/subspecies of a crop. We believe that our empirical specification, explained in section five, partly addresses some of these concerns.

3.4 Defining cereal diversity

Quantifying diversity is complicated and, so far, no universal definition has been agreed upon. A number of different definitions have been proposed (Baumgärtner, 2006) but different definitions are used in different contexts, not least because different professions value biodiversity for different reasons (see Baumgärtner (2006) for a review of the debate). For our purpose, the most common indices used include a simple count measure (used in Di Falco et al., 2010),

[1]Gebregziabher et al. (2012) find that, in the Tigray region, the yield response to chemical fertilizer is poor under rain-fed conditions since it is a moisture-strained environment.

the Simpson index and the Shannon Index (used by Di Falco and Chavas, 2008). In this paper, we opt for the use of the Shannon Index of cereal diversity for three reasons.

First, as argued by Di Falco and Chavas (2008), it is possible that a simple index of species richness, which fails to control for evenness, will lead to a "sampling effect". As a result, the diversity index may capture the performance of a single species (crops in our case) rather than the effect of diversity. However, since the Shannon index controls for both richness and evenness this problem becomes less severe.

Secondly, the Shannon index is likely to be more suitable than the Simpson index in this context, as it has been found in the literature that the Simpson index could be biased towards the dominant species (Magurran, 1988; Di Falco and Chavas, 2008).

Thirdly, data for constructing alternative indices of cereal diversity were not available to build the index proposed by Weitzman (1992), which would be suitable. The index proposed by Weitzman (1992) is a measure of genetic distance. However, the data required for the construction of such an index is simply not available in this dataset.

There are three limitations of the Shannon index in this application. First, while we observe different cereals, we do not observe different sub-species of the same crop[2]. This was shown to be important in a number of studies, including Di Falco and Chavas (2008), Di Falco et al. (2007) and Di Falco and Chavas (2006). This is an issue we are not able to address given our data. We only focus on crop diversity (across different crops), as was done in Di Falco et al. (2010), Bangwayo-Skeete et al. (2012) and Perrings and Di Falco (2003). A second limitation is that our Shannon index is built at the household level. As a result, it is possible that, in some cases, a non-negative Shannon index captures two monocultures in separate plots[3]. Finally, a third limitation is that we look at the Shannon index for cereals only. Cereals has been the most common focus of this literature, whether exploring crop-specific diversity, cereal-specific diversity or through count indices where cereals account for the majority of the crops cultivated. However, measuring other types of crop diversity could potentially lead to different effects on crop production.

[2] With the exception of teff, where we observe both white and black teff.

[3] However, in our data we do not have information about the location of different plots. As such, while it could be that the two monocultures are in plots very far away from each other, it could also be that they are located in plots adjacent to each other. As such, it is not clear whether building the Shannon index at the plot level would be preferable.

As in Di Falco and Chavas 2008, we calculate the cereal Shannon index as follows:

$$SI = -\sum_{i} p_i * log(p_i) \quad (3.1)$$

Where p_i represents the proportion (of cereal area) allocated to cereal crop i. Given that we include six cereals in the analysis, the Shannon index has a theoretical range between 0-1.8[4].

3.5 Data and methodology

3.5.1 Data

The dataset used is the Ethiopian Rural Household Survey (ERHS 2011)[5] and all waves since 1994 are used. The 1994 wave is composed of 1,470 households from 18 different peasant associations (15 different villages), spread over four regions. The location, characteristics and the agro-ecological zone breakdown of these peasant associations can be found in figures A1 and A2 and Table A1, respectively (Appendix A)[6]. However, it is important to mention that this sample is not nationally representative (Dercon and Hoddinott, 2004).

As mentioned in Dercon and Hoddinnot (2004), attrition between 1994-2004 is estimated at 13%. In addition, only observations that cultivate cereals in at least two consecutive periods were used in our sample. This choice was driven by the needs of the semi-parametric model. As a result, the sample in this paper consists of 1280 individuals (5804 observations), for which a table of summary statistics (Table 3.1) is presented below.

[4] A household cultivating two cereals in equal proportions will have a Shannon index of 0.69. If three cereals are cultivated in equal proportions, the Shannon index will take a value of approximately 1.1.

[5] These 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.

[6] The agro-ecological zone breakdown was adapted from Hoddinott et al. 2011. Dercon and Hoddinott 2004 is the source for the map and site characteristics.

Table 3.1: Summary Statistics

	All		N. Highlands		C. Highlands		Other		Enset	
	Mean	**S.D.**	**Mean**	**S.D.**	**Mean**	**S.D.**	**Mean**	**S.D.**	**Mean**	**S.D.**
Cereal Production (kgs)	810.12	988.32	343.69	414.28	1073.79	946.29	1260.72	1300.63	256.18	385.94
Cereal Yield (kg/ha)	806.22	753.13	518.14	518.58	922.89	766.19	972.23	799.23	714.35	795.18
Cereal Area (ha)	1.20	1.09	0.87	0.95	1.43	1.01	1.64	1.25	0.53	0.65
Shannon index	0.49	0.41	0.45	0.43	0.50	0.34	0.77	0.37	0.13	0.26
Number of oxen	0.87	1.10	0.64	0.83	1.20	1.13	0.95	1.28	0.41	0.84
Household Size	6.02	2.72	5.21	2.39	5.84	2.65	6.30	2.58	7.04	3.04
Quantity Ferilizer (kgs)	51.38	85.51	3.03	11.28	77.65	85.19	80.45	117.92	21.13	29.15
Number of ploughs (units)	1.78	2.98	1.76	3.05	2.31	3.32	1.66	2.99	0.94	1.69
Number of hoes (units)	1.12	1.59	0.82	1.40	1.41	1.80	1.07	1.56	1.00	1.30
Tigray	0.13	0.33	0.56	0.50	0.00	0.00	0.00	0.00	0.00	0.00
Amhara	0.37	0.48	0.44	0.50	0.78	0.42	0.00	0.00	0.00	0.00
Oromia	0.33	0.47	0.00	0.00	0.22	0.42	1.00	0.00	0.00	0.00
SSN	0.18	0.38	0.00	0.00	0.00	0.00	0.00	0.00	1.00	0.00
Northern Highlands	0.23	0.42	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Central Highlands	0.34	0.48	0.00	0.00	1.00	0.00	0.00	0.00	0.00	0.00
Other	0.25	0.43	0.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00
Enset	0.18	0.38	0.00	0.00	0.00	0.00	0.00	0.00	1.00	0.00
Number of observations	5804		1323		2002		1456		1023	

Table 3.1 highlights stark differences in terms of the use of inputs across different agro-ecological zones. Overall, farmers in the Central Highlands and in the Arusi/Bale ("Other") agro-ecological zones allocate higher proportions of land to cereals, use more fertilizer, have higher average levels of cereal diversity and display the highest yields compared to the average household in the Northern Highlands or in the Enset agro-ecological zones.

In terms of the variables used in this paper, the dependent variable in this study is the total production of cereals, which sums the production (in kilograms) of each cultivated cereal. The explanatory variables included consist of cereal area (measured in ha), number of oxen, household size (to proxy for labour), the quantity of fertilizer[7], the number of hoes and ploughs. In addition to this, the crop diversity variable, the cereal Shannon Index, will be included. A detailed explanation of how these variables were constructed is available in Appendix B.

[7]In the case of fertilizer, whenever there was data on the application of fertilizer directly on cereal, this data was used. When only the total amount of fertilizer was available, the total amount was apportioned to cereal area (i.e. we assumed the households used fertilizer evenly in their land).

3.5.2 Methodology

Fixed Effects model

Our analysis of the productive effects of crop diversity is concerned with agricultural productivity and the role of crop diversity.

Concerning the functional form, we opt for a translog functional form, which includes the natural logarithm of land, labour, fertilizer, oxen, hoes, crop diversity (Shannon index) their squares and their interactions. In order to capture local level trends in output as well as aspects such as weather shocks which are common to households in a given peasant association, we also include a dummy variable for each peasant association-year[8]. We prefer to include peasant-association-year dummy variables rather than a time trend since we do not want to impose a specific time trend at the peasant association level.

We estimate a fixed effects model as it accounts for unobserved heterogeneity at the household level. The estimated regression can be algebraically expressed as follows:

$$ln\, y_{it} = \alpha_i + \sum_{n=1}^{n=N} \beta_n ln(x_{nit}) + 0.5 \sum_{n=1}^{n=N} \sum_{m=1}^{m=N} \beta_{nm}\, lnx_{nit} * lnx_{mit} + \sum_{t=1}^{t=T} \sum_{p=1}^{p=P} d_t * d_p + e_{it} \quad (3.2)$$

Equation 3.2 can be interpreted in four parts. First, α_i captures household-specific, time-invariant features. The second part refers to the inclusion of the natural logarithms of all the explanatory variables, their squares and their interactions. The third part of this equation refers to the year-peasant association dummy variables ($d_t * d_p$), which absorb common shocks at the peasant association level for different years[9]. Finally, the last component is the error term, e_{it}. We note that a number of our variables (oxen, fertilizer) have a large proportion of 0 values. This has been shown to potentially lead to biased estimates of the marginal effects if not dealt with properly (Battese, 1997). In our case, however, there is often little within variation in input-use[10] which could make this approach slightly problematic when

[8] i.e. for each peasant association we include a dummy for each year.

[9] This is likely to include aspects such as rainfall, temperature, as well as peasant-association specific trends in production over time.

[10] i.e. many farmers either use or do not use a given input in most periods, although the quantities do

using fixed-effects. Nevertheless, we show (in Appendix A, Tables 3A.11-3A.16) that using the correction proposed by Battese (1997) does not alter the main conclusions of the paper, though the magnitudes become quite different. We also briefly explain the correction proposed by Battese (1997), its rationale and how it is implemented in practice in Appendix B.

3.5.3 Semiparametric regression estimator

In addition to the parametric results, since there is little theoretical guidance on the likely shape of the relationship between cereal diversity and production, we also conduct a series of semi-parametric regressions. This specification allows for greater flexibility in the relationship, since it makes it easier to investigate a possible non-linear effect of crop diversity on production. The basic cross section model proposed by Robinson (1988) can be summarized using the following equation:

$$ln\, y_{it} = \boldsymbol{\beta}_x \mathbf{X}_{it} + f(sh_{it}) + e_{it} \tag{3.3}$$

Where $\mathbf{X}$ is a set of explanatory variables which includes all inputs except the Shannon index. For the parametric component of the model, we use the two sets of variables detailed in the previous section, but exclude the Shannon index. The latter is captured in the component $f(sh_{it})$, which represents the non-parametric smooth function of the Shannon index, which we believe may be non-linear.

This model has subsequently been extended to include fixed effects in a panel data setting (Baltagi and Li, 2002). The Baltagi and Li (2002) model differs from the original model by taking the first differences of equation 3.3. We implement this procedure using the *xtsemipar* command in STATA 14 (Libois and Verardi, 2012). For all sets of results, we use a kernel regression with the rule-of-thumb bandwidth. In all cases, we use a degree 4 local weighted polynomial fit using the Epanechnikov kernel[11].

change

[11] We also test the sensitivity of our results to a degree one local polynomial fit.

3.5.4 Limitations of the empirical approach

The first and most important limitation of our approach, as with other papers in this literature, relates to the issue of endogeneity. Given that the choice of the level of cereal diversity is likely to be endogenous and that we were unable to find a suitable instrument, we are not able to claim the estimation of a causal relationship between cereal diversity and production. However, our empirical specification is more stringent than the norm in similar studies, thereby potentially attenuating concerns related to endogeneity. Specifically, we take two steps that make for a more convincing approach to the estimation of this relationship than what has traditionally been the case in the literature. First, we use panel data and, as a result, this allows us to control for household fixed effects, which are likely to control for household-specific, time-invariant characteristics. Secondly, all of our specifications use peasant association-year fixed effects, which are likely to control for common, time-varying unobservable heterogeneity at the peasant-association level.

A second limitation relates to the narrow focus of our question as we focus solely on the effect of cereal diversity on cereal production. This has been the most common approach in agricultural economics. However, it is possible that other types of diversity (such as mixing a cereal with a legume, for instance) may have a very different effect on production.

Finally, we focus only on the productive implications of the diversity of systems of cereal production. We do not discuss the relationship between crop diversity and other production or environmental variables, such as volatility or erosion, which may be important and pertinent.

3.6 Results

3.6.1 Parametric results

As can be seen from Table 3.2, the estimated coefficients associated with the Shannon index differ substantially from one agro-ecological zone to another. Concerning the overall productive effect of cereal diversity on production, as has been the norm in the literature, we find a positive and statistically significant effect between cereal diversity and cereal production for the full sample (column 1, Table 3.2). However, running the regressions separately by agro-ecological zone reveals very stark differences. Although we find a positive elasticity in every agro-ecological zone (columns 2-5, Table 3.2), this elasticity is only large and statistically significant in the "Other agro-ecological zones (Arusi/Bale). An interesting aspect from these results is that the ordering of the magnitude mirrors closely the proportion of households who cultivate teff, which is known to be a lower productivity crop. In other words, the agro-ecological zone displaying the highest coefficients is also the agro-ecological zone where teff is most prevalent.

Table 3.2: Main results : Parametric translog

	(1) All	**(2)** N. Highlands	**(3)** C. Highlands	**(4)** Other	**(5)** Enset
Shannon index	0.046	0.026	0.030	0.122	0.032
	(0.029)	(0.054)	(0.047)	(0.075)	(0.068)
Shannon index (square)	0.003	0.003	0.002	0.007	0.001
	(0.002)	(0.004)	(0.003)	(0.005)	(0.004)
Area*Shannon index	-0.006**	-0.009	-0.006	0.008	-0.017**
	(0.003)	(0.005)	(0.004)	(0.008)	(0.007)
Household size*Shannon index	0.003	0.008	0.000	-0.006	-0.007
	(0.004)	(0.007)	(0.006)	(0.011)	(0.011)
Oxen*Shannon index	0.000	0.000	0.000	-0.001	0.000
	(0.000)	(0.001)	(0.000)	(0.001)	(0.001)
	(0.000)	(0.001)	(0.000)	(0.000)	(0.001)
Fertilizer*Shannon index	0.000	0.000	0.000	0.000	-0.001**
	(0.000)	(0.001)	(0.000)	(0.001)	(0.001)
Hoes*Shannon index	0.000	0.001***	0.000	0.000	-0.002**
	(0.000)	(0.000)	(0.000)	(0.001)	(0.001)
Ploughs*Shannon index	0.000	-0.001*	0.000	-0.001	0.002***
	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)
Constant	6.339***	4.919***	6.450***	5.943***	5.440***
	(0.110)	(0.353)	(0.171)	(0.234)	(0.411)
Elasticity of Shannon index	0.044*	0.035	0.025	0.115*	0.031
p-value	0.061	0.401	0.529	0.084	0.432
Fixed effects	✓	✓	✓	✓	✓
Village-year fixed effects	✓	✓	✓	✓	✓
Linear variables	✓	✓	✓	✓	✓
Squares	✓	✓	✓	✓	✓
Interactions	✓	✓	✓	✓	✓
Number of households	1280	289	428	299	264
Number of observations	5804	1323	2002	1456	1023
Average obs. per household	4.534	4.578	4.678	4.87	3.875
R-squared a	0.547	0.659	0.557	0.509	0.454
R-squared w	0.556	0.672	0.571	0.526	0.481

Notes: N. Highlands refers to Northern Highlands. C. Highlands refers to Central highlands. Numbers in parentheses represent clustered standard errors at the household level. The specification in the regression is a translog specification and the full list of coefficients can be seen in Table 3A.2 in the Appendix. As explained in the methodology section, this specification does not include the adjustment proposed by Battese (1997) since there is little within-household variation of input-use. Instead 0 values are assigned the value of 0.000001.

We thus test whether the effect we capture could be attributed to the cultivation of teff and break the sample into households that cultivate teff and those who do not. These results can be seen in Tables 3.3-3.4. Overall, the results in Table 3.3, which only include households who cultivate teff, seem to suggest a positive significant elasticity of cereal diversity in two out of four agro-ecological zones. However, once households that cultivate teff are removed, none of the elasticities are statistically significant, though in one case the coefficient increases. These results do not prove beyond doubt that the full effect is attributable to a compositional effect. For one, sample sizes decrease substantially in a number of agro-ecological zones, which makes it harder to ascertain statistical significance. Nevertheless, these results are indicative that a compositional effect could partly explain the relationship. Perhaps what is being captured in these results is that, as cereal diversity increases, the relative contribution of the low productivity cereal gradually fades, thereby leading to higher levels of production and productivity.

Table 3.3: Main results : Parametric translog (teff only)

	(1)	**(2)**	**(3)**	**(4)**	**(5)**
	All	N. Highlands	C. Highlands	Other	Enset
Shannon index	0.177***	0.129	0.065	0.299***	0.190
	(0.047)	(0.086)	(0.066)	(0.076)	(0.133)
Shannon index (square)	0.011***	0.015**	0.003	0.018***	0.006
	(0.003)	(0.006)	(0.004)	(0.005)	(0.009)
Area*Shannon index	-0.019***	-0.035***	-0.017**	-0.009	-0.022**
	(0.005)	(0.012)	(0.009)	(0.016)	(0.009)
Household size*Shannon index	0.000	0.028	-0.006	-0.015	-0.040***
	(0.006)	(0.018)	(0.011)	(0.015)	(0.015)
Oxen*Shannon index	0.000	-0.001	0.000	0.001	0.001
	(0.000)	(0.002)	(0.001)	(0.001)	(0.001)
Fertilizer*Shannon index	0.000	-0.002	0.000	0.000	-0.001
	(0.000)	(0.002)	(0.001)	(0.001)	(0.001)
Hoes*Shannon index	0.001	0.002	0.002**	0.002*	-0.002*
	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)
Ploughs*Shannon index	0.001**	-0.002	0.001	-0.003**	0.003***
	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)
Constant	6.345***	5.469***	6.846***	5.786***	5.771***
	(0.146)	(0.553)	(0.340)	(0.232)	(0.563)
Elasticity of Shannon index	0.153***	0.186**	0.036	0.252***	0.096
p-value	0.000	0.019	0.536	0.000	0.254
Fixed effects	✓	✓	✓	✓	✓
Village-year fixed effects	✓	✓	✓	✓	✓
Linear variables	✓	✓	✓	✓	✓
Squares	✓	✓	✓	✓	✓
Interactions	✓	✓	✓	✓	✓
Number of households	782	152	217	211	202
Number of observations	2799	544	679	960	616
Average obs. per household	3.579	3.579	3.129	4.55	3.05
R-squared a	0.511	0.358	0.557	0.597	0.535
R-squared w	0.526	0.412	0.59	0.616	0.573

Notes: N. Highlands refers to Northern Highlands. C. Highlands refers to Central highlands. Numbers in parentheses represent clustered standard errors at the household level. The specification in the regression is a translog specification and the full list of coefficients can be seen in Table 3A.3 in the Appendix. As explained in the methodology section, this specification does not include the adjustment proposed by Battese (1997) since there is little within-household variation of input-use. Instead 0 values are assigned the value of 0.000001.

Table 3.4: Main results : Parametric translog (no teff)

	(1) All	**(2)** N. Highlands	**(3)** C. Highlands	**(4)** Other	**(5)** Enset
Shannon index	-0.004	-0.094	0.095	-0.069	0.011
	(0.050)	(0.072)	(0.081)	(0.136)	(0.131)
Shannon index (square)	-0.001	-0.007	0.006	-0.009	-0.002
	(0.003)	(0.005)	(0.006)	(0.009)	(0.008)
Area*Shannon index	-0.002	0.004	-0.002	0.012	-0.017
	(0.004)	(0.009)	(0.005)	(0.011)	(0.026)
Household size*Shannon index	0.000	0.005	-0.004	-0.029*	-0.016
	(0.005)	(0.008)	(0.007)	(0.016)	(0.035)
Oxen*Shannon index	0.000	0.000	0.000	-0.001	0.000
	(0.000)	(0.001)	(0.001)	(0.001)	(0.002)
Fertilizer*Shannon index	0.000	0.000	0.000	0.001	0.000
	(0.000)	(0.001)	(0.000)	(0.001)	(0.002)
Hoes*Shannon index	0.000	0.001	-0.001	0.000	0.001
	(0.000)	(0.001)	(0.001)	(0.001)	(0.002)
Ploughs*Shannon index	0.000	0.000	0.000	0.000	0.003
	(0.000)	(0.001)	(0.001)	(0.001)	(0.003)
Constant	6.305***	4.852***	6.347***	6.115***	5.745***
	(0.175)	(0.537)	(0.222)	(0.728)	(1.268)
Elasticity of Shannon index	0.000	-0.058	0.076	-0.105	-0.016
p-value	0.997	0.219	0.272	0.382	0.838
Fixed effects	✓	✓	✓	✓	✓
Village-year fixed effects	✓	✓	✓	✓	✓
Linear variables	✓	✓	✓	✓	✓
Squares	✓	✓	✓	✓	✓
Interactions	✓	✓	✓	✓	✓
Number of households	893	211	344	128	210
Number of observations	3005	779	1323	496	407
Average obs. per household	3.365	3.692	3.846	3.875	1.938
R-squared a	0.547	0.659	0.557	0.509	0.454
R-squared w	0.556	0.672	0.571	0.526	0.481

Notes: N. Highlands refers to Northern Highlands. C. Highlands refers to Central highlands. Numbers in parentheses represent clustered standard errors at the household level. The specification in the regression is a translog specification and the full list of coefficients can be seen in Table 3A.4 in the Appendix. As explained in the methodology section, this specification does not include the adjustment proposed by Battese (1997) since there is little within-household variation of input-use. Instead 0 values are assigned the value of 0.000001.

Whether, if it exists, this compositional effect can still be reconciled with the traditional channels through which crop diversity can lead to over-yielding is difficult to answer. On the one hand, the existence of such a compositional effect would probably be at odds with the "complementarity" and the "facilitation" effects since cereals are likely to be quite similar and may not differ a lot in the timing of resources. However, it could be argued that it could represent a conscious "sampling" effect since the household may well select the crop that performs the best (most productive), given the environment. However, if the composition of cereals is driving the result, this questions the extent to which promoting cereal diversity could lead to improved agricultural production. The reason being that, if a compositional effect is driving the result, this implies that increases in diversity only lead to increases in productivity when farmers add a high(er) productivity crop to their crop mix. Increases in diversity in the other direction would not yield increases in productivity. From a policy perspective, therefore, if this effect is the main driver of productivity increases, the policy implication is that systems should promote more productive crops.

We also carry out three sets of robustness checks, which are available in Appendix A. Tables 3A.11-3A.16 summarize the results when the regression is estimated by OLS and the Battese (1997) correction is applied. Tables 3A.18-3A.23 show the results when we only consider households for which there is no imputed data for fertilizer, ploughs and hoes. Finally, Tables 3A.31-3A.36 summarize the results when only the households for which we have six observations are considered (i.e. a balanced panel)[12]. Though magnitudes certainly differ, the overarching conclusions remain the same in all three robustness checks.

3.6.2 Semi-parametric results

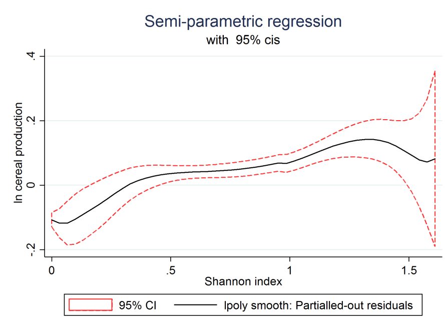
Given that there is no proven underlying theory informing the expected shape and magnitude of the production-diversity relationship, the statistically insignificant results displayed in the previous section may be masking existing non-linearities. In other words, it is possible that the insignificant result in parametric models are a result of not taking into account non-linearities appropriately. Alternatively, it is possible that a positive effect of crop diversity

[12]Balanced subsample refers to the sub-sample of households for which we have observations for each period. However, the teff only and no teff regression are not necessarily balanced since some households switch in and out of teff during the sample period. Also, as can be seen in Tables 3A.31-3A.33, using the balanced sample leads to a sharp decrease in sample size. This is very severe in the Enset area for the no-teff subsample, where there are very few observations with a Shannon index above 0. As a result, for this subsample, we do not have a high degree of confidence in the results presented.

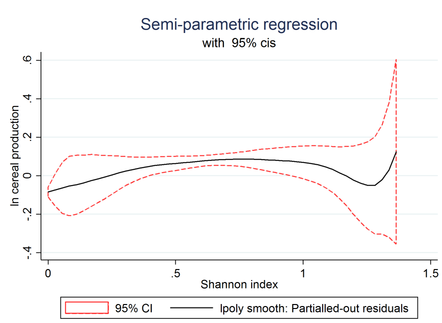
exists, but that this effect is confined to a subset of the Shannon index values. Additionally, it is also possible that a relationship exists but that statistical significance is hampered by the small sample size of some of these subsamples. It is for these reasons that we also run a set of semi-parametric regressions, which allow for a more flexible characterization of the relationship between crop diversity and production and tend to be less sensitive to sample size.

The parametric part of the results is summarized in Tables 3A.5-3A.9 and the smooth functions of the crop diversity result on the partialled-out residuals are available in Figures 3.1-3.5[13]. The same local polynomial including the scatter plots are also available in the Appendix (figures 3A.3-3A.7). For each figure corresponding to a given geographical region, three subfigures are presented. Subfigure (a) summarizes the results when all the households in a given region are included, subfigure (b) summarizes the results when teff producers are excluded and subfigure (c) shows the results when only teff producers are included.

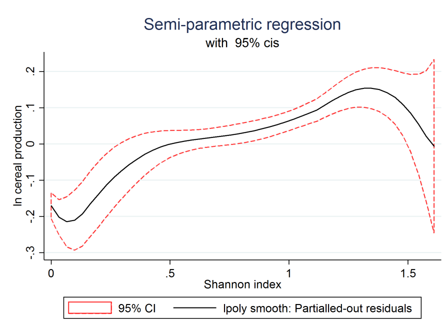
Figure 3.1: Effect of Shannon index Semi parametric Full sample



(a) Full sample



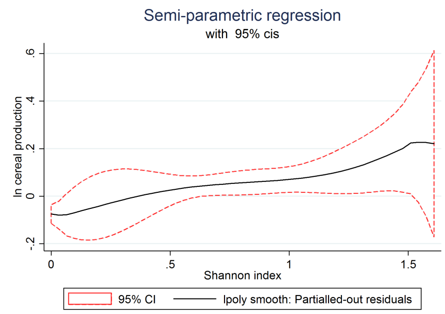
(b) Non teff-producing households



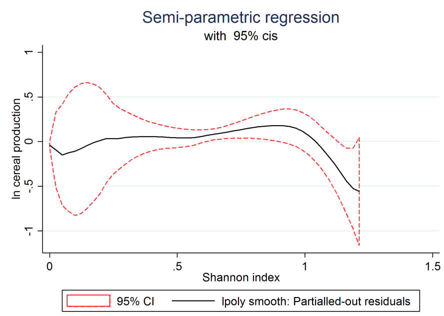
(c) Teff-producing households

[13]Figures which include the scatter plots are available in the Appendix Tables 3A.3-3A.7. Both sets of figures (3.1-3.5 and 3A.3-3A.7) use a degree 4 polynomial (the default) and the rule-of-thumb bandwidth. The rule-of-thumb bandwidth is summarized in the Appendix in Table 3A.10.

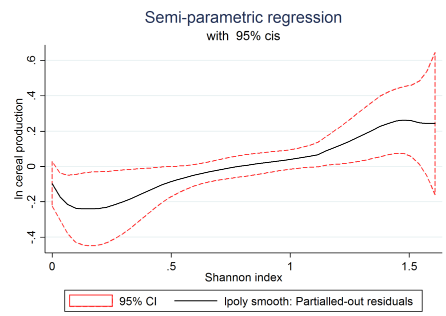
Figure 3.2: Effect of Shannon index Semi parametric Northern Highlands



(a) Full sample

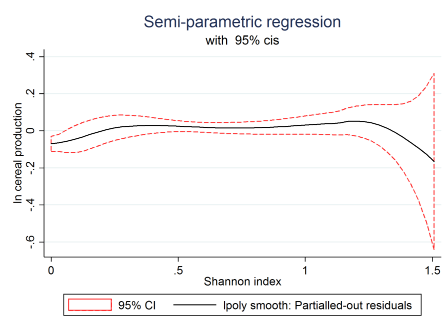


(b) Non teff-producing households

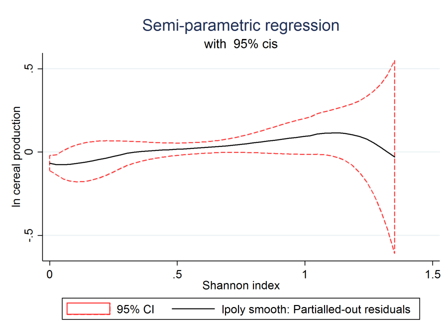


(c) Teff-producing households

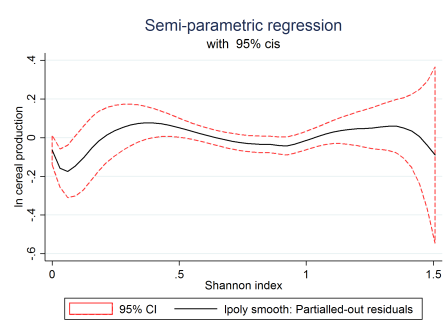
Figure 3.3: Effect of Shannon index Semi parametric Central Highlands



(a) Full sample

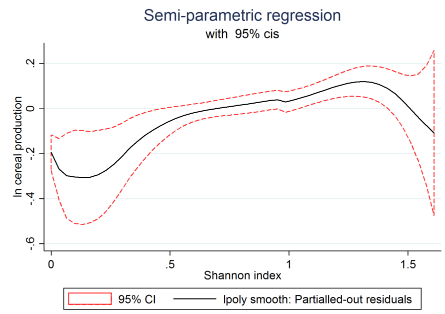


(b) Non teff-producing households

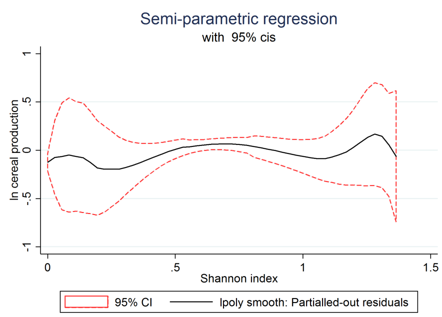


(c) Teff-producing households

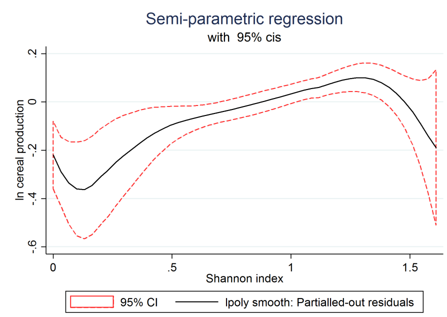
Figure 3.4: Effect of Shannon index Semi parametric Arussi/Bale



(a) Full sample

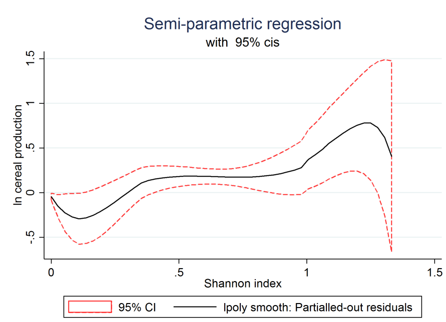


(b) Non teff-producing households

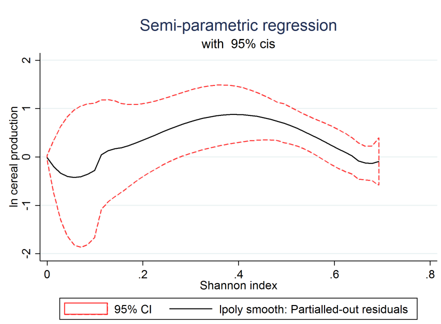


(c) Teff-producing households

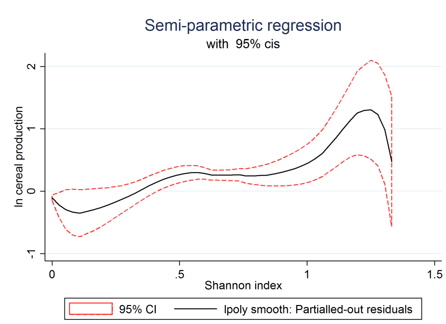
Figure 3.5: Effect of Shannon index Semi parametric Enset



(a) Full sample



(b) Non teff-producing households



(c) Teff-producing households

The semi-parametric results, to a certain extent, confirm the findings of the parametric results. We find a clear positive correlation between the Shannon index for panel (a) of the full sample (Figure 3.1) and in the Arusi/Bale/Hararghe agro-ecological zones (Figure 3.4), with the Northern Highlands (Figure 3.2) also displaying a positive, but somewhat noisy relationship. Also similar to our findings from the parametric models, these results appear to be largely driven by the inclusion of teff producers, with none of the panels (b) displaying a large, clear and positive relationship, though panel (b) of Figure 3.4 seems to suggest a somewhat positive relationship[14]. Conversely, most panels (c), with the exception of panel (c) in Figure 3.3, suggest a positive relationship, indicating that the inclusion of teff producers seems to partly drive our results. This provides some support for the existence of a potential compositional effect.

However, the semiparametric results also shed some light on additional aspects of this relationship. First, when we find a clear positive relationship, we tend to also find a large, statistically significant negative intercept. For the Arusi/Bale/Hararghe "Other") agro-ecological zones, this can be explained by the fact that the vast majority of farmers who cultivate one crop cultivate teff. As a result, an increase in the crop diversity index could indicate a shift away from a low productivity cereal. A similar mechanism may be at play in the Northern Highlands for the case of barley, which is the second lowest productivity crop in our sample. Second, the semi-parametric results also highlight aspects related to the shape of the relationship. Partly as a result of the negative intercept, in some cases (Figures 3.2, 3.3 and 3.5), the strongest positive relationship occurs between low to medium values of the Shannon index. We also find a sharp (but very noisy) decrease in the relationship at large levels of cereal diversity in three out of four agro-ecological zones (Figure 3.2 being the exception). Taken together, these results suggest rather limited benefits of pursuing very diverse systems in terms of cereal production.

[14]Panel (b) in Figure 3.5 is not particularly informative as there are very few observations of non-teff producer with a non-0 Shannon index (less than 5% of the values). This is made more clear in Figure 3A.7.

3.7 Conclusion

This paper revisited the link between cereal diversity and cereal production using a panel representative of a larger geographical area in Ethiopia than what has typically been the case in the literature. Doing this allows us to understand whether *in situ* conservation may yet deliver a promising solution in terms of conservation of plant genetic material alongside sustained productivity gains.

In some cases, our results corroborate a number of previous results in the literature. For instance, we find large positive gains of cereal diversity on cereal production for the full sample. However, we also find that these effects are very heterogeneous across agro-ecological zones. Specifically, certain agro-ecological zones (Arusi/Bale/Hararghe) and one crop (teff) seem to be driving these results in both parametric and semi-parametric specifications. This suggests that, at least in our case, the "biodiversity" effect seems to be capturing a decreasing share of a low productivity crop in the crop-mix.

Whether this can be reconciled with the typical channels used to explain the crop diversity ("biodiversity")-productivity link is arguable. The fact that this result seems to be driven by teff suggests that this result is at odds with the "complementarity" and the "facilitation" effect. However, in a way this could be considered a deliberate "sampling" effect.

These results highlight the importance for practitioners in the literature to attempt to understand what is driving the results between diversity and productivity. It is important to at least consider the possibility that this effect could partly reflect different potential yields for cereals in the crop mix. As a result, increases in the diversity index could be capturing a move away or towards a particularly high- or low-performing cereal. In our case, given that we do not have data on subspecies, the results seem to be partly driven by one crop (teff). However, a similar mechanism could be at play with particular high- or low-performing subspecies of a given crop.

From a policy perspective, however, the results highlight two main points. Firstly, while diversity, in itself, may be desirable for a number of reasons, its positive productive implications are not clear once farmers who cultivate low-yield crops are removed from the sample. As a result, it seems that increases in diversity only seem to have a positive effect in one direction (from high proportion of low-yield crop to diverse mix of low- and high-yield crops). Secondly, the shape revealed in the semi-parametric method suggests that these effects are not linear and that, beyond a certain point, the associated gains of increased diversity seem tenuous, at best.

Taken together, these results suggest that cereal diversity is unlikely to be a panacea for cereal productivity. Lack of clear production gains from increasing cereal diversity allied to the development of alternative sources of insurance and the modernization of agriculture, which tends to lead to a reduction of cereal diversity, highlights the need to focus on alternative means of conserving crop genetic diversity.

In addition to this, our paper highlights a number of possible directions for future research. First, this paper focuses on a very narrow type of crop diversity (cereal diversity) and these results are not necessarily transposable to other types of crop diversity, for which the relationship may be very different. Second, while we believe our empirical specification improves on previous literature focusing on this question, endogeneity remains a concern. Consequently, our results do not settle this debate and we cannot and do not claim a causal relationship. Further research regarding potential instruments or alternative research designs (e.g. field experiments) would be useful. A third aspect that was absent from this analysis relates to the relationship between land degradation and crop diversity. As argued by Taddese (2001), land degradation is a serious issue in Ethiopia and crop diversity may well have an important effect on land quality, which we do not capture or investigate in this paper. Finally, our analysis leaves aside the links between cereal diversity and income, nutrition as well as production and income volatility, all of which could be valid reasons to pursue a diversification strategy, despite limited gains in output. In our specific case, while teff typically displays lower yields, it has a very high market value compared to other cereals. As a result, it could still make economic sense to cultivate teff, despite its productive implications.

References

[1] Baltagi, B. H. & Li, D. (2002). Series estimation of partially linear panel data models with fixed effects. *Ann. Econ. Finance 3*, 103116.

[2] Bangwayo-Skeete, P. F., Bezabih, M., & Zikhali, P. (2012). Crop biodiversity, productivity and production risk: Panel data microevidence from Ethiopia. *Natural Resources Forum*, 36(4), 263-273.

[3] Baumgärtner, S. (2006). Measuring the diversity of what? And for what purpose? A conceptual comparison of ecological and economic biodiversity indices. And for What Purpose? A conceptual comparison of ecological and economic biodiversity indices.

[4] Baumgärtner, S. (2007). The insurance value of biodiversity in the provision of ecosystem services, *Natural Resource Modelling* 20(1), 87-127.

[5] Baumgärtner, S. & M.F. Quaas (2009), Agro-biodiversity as natural insurance and the development of financial insurance markets, in: A. Kontoleon, U. Pascual and M. Smale (eds), Agrobiodiversity, Conservation and Economic Development, Routledge, London, pp. 293-317.

[6] Baumgärtner, S., & Quaas, M. F. (2010). Managing increasing environmental risks through agrobiodiversity and agri-environmental policies. *Agricultural Economics*, 41(5), 483-496.

[7] Battese, G. E. (1997). A note on the estimation of CobbDouglas production functions when some explanatory variables have zero values. *Journal of agricultural Economics*, 48(13), 250-252.

[8] Benson, T., Ermias, E. & Thurlow, J., 2014. The economywide effects of teff, wheat,

and maize production increases in Ethiopia: Results of economywide modeling. *IFPRI Discussion Paper 1366*

[9] Borlaug, N. (2000). The Green revolution revisited and the road ahead. Special 30th Anniversary Lecture, The Norwegian Nobel Institute, Oslo, September 8, 2000.

[10] Collier, P., & Dercon, S. (2014). African Agriculture in 50 Years: Smallholders in a Rapidly Changing World? *World Development*, 63, 92-101.

[11] Dercon, S., & Zeitlin, A. (2009). Rethinking agriculture and growth in Ethiopia: A conceptual discussion. Review of the Agricultural Strategy in Ethiopia

[12] Chavas, J. P., & Di Falco, S. (2012a). On the productive value of crop biodiversity: evidence from the highlands of Ethiopia. *Land Economics*, 88(1), 58-74.

[13] Chavas, J. P., & Di Falco, S. (2012b). On the role of risk versus economies of scope in farm diversification with an application to Ethiopian farms. *Journal of agricultural economics*, 63(1), 25-55.

[14] Dercon, S., & Hoddinott, J. (2004). The Ethiopian rural household surveys: Introduction. International Food Policy Research Institute, Washington, DC.

[15] Di Falco, S. (2012). On the value of agricultural biodiversity *Annual Review of Resource Economics*, 4, 207-223.

[16] Di Falco, S., Bezabih, M., & Yesuf, M. (2010). Seeds for livelihood: Crop biodiversity and food production in Ethiopia. *Ecological Economics*, 69(8), 1695-1702.

[17] Di Falco, S. & Chavas, J-P. (2006). Crop genetic diversity, farm productivity and the management of environmental risk in rainfed agriculture. *European Review of Agricultural Economics*, 33(3), 289-314.

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

[19] Di Falco, S., & 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.

[20] Di Falco, S., Chavas, J-P. & Smale, M. (2007). Farmer management of production risk on degraded lands: the role of wheat variety diversity in the Tigray region, Ethiopia.

Agricultural Economics, 36(2), 147-56.

[21] Di Falco, S. & Perrings, C. (2003). Crop genetic diversity, productivity and stability of agro-ecosystems: A theoretical and empirical investigation *Scottish Journal of Political Economy*, 50(2), 207-16.

[22] Di Falco, S., & Perrings, C. (2005). Crop biodiversity, risk management and the implications of agricultural assistance. *Ecological economics*, 55(4), 459-466.

[23] Ethiopia Rural Household Survey Dataset, 1989-2009. (2011). Washington, D.C. International Food Policy and Research Institute (IFPRI) (datasets).

[24] Evenson, R. E., & Gollin, D. (2003). Assessing the impact of the Green Revolution, 1960 to 2000. *Science*, 300(5620), 758-762.

[25] Gebregziabher, G., & Holden, S. (2011). Does irrigation enhance and food deficits discourage fertilizer adoption in a risky environment? Evidence from Tigray, Ethiopia. *Journal of Development and Agricultural Economics*, 3(10), 514-528.

[26] Hooper, D. U., Chapin, F. S., Ewel, J. J., Hector, A., Inchausti, P., Lavorel, S., Lawton, J. H., Lodge, D. M., Loreau, M., Naeem, S., Schmid, B., Setala, H., Symstad, A. J., Vandermeer, J. & Wardle, D. A. (2005). Effects of biodiversity on ecosystem functioning: a consensus of current knowledge. *Ecological monographs*, 75(1), 3-35.

[27] Libois, F. & Verardi, V. 2012. Semiparametric fixed-effects estimator. *Working Paper*. University of Namur.

[28] Nisrane, F., Guush, B., Sinafikeh, A., Gerawork, G., Taffesse, A. S. & Hoddinott, J. (2011). Sources of Inefficiency and Growth in Agricultural Output in subsistence Agriculture: A stochastic Frontier Analysis. *International Food Policy and Research Institute, ESSP II Working Paper*, (19).

[29] Omer, A., Pascual, U., & Russell, N. P. (2007). Biodiversity conservation and productivity in intensive agricultural systems. *Journal of Agricultural Economics*, 58(2), 308-329.

[30] Quaas, M.F. & Baumgärtner, S. (2008), Natural vs. financial insurance in the management of public good ecosystems, Ecological Economics 65(2), 397-406.

[31] Rahmato, D. (2008). Ethiopia: agriculture policy review. Digest of Ethiopias National

Policies, Strategies, and Programs.

[32] Robinson, P.M. (1988). Root-N-consistent semiparametric regression. *Econometrica*, 56(4), 931954.

[33] Schlenker, W. & Lobell, D.B., 2010. Robust negative impacts of climate change on African agriculture. *Environmental Research Letters*, 5(1), p.014010.

[34] Smale, M., Hartell, J., Heisey, P. W., and Senauer, B. (1998). The contribution of genetic resources and diversity to wheat production in the Punjab of Pakistan. *American Journal of Agricultural Economics*, 80(3), 482-493.

[35] Smale, M., Singh, J., Di Falco, S., & Zambrano, P. (2008). Wheat breeding, productivity and slow variety change: evidence from the Punjab of India after the Green Revolution. *Australian Journal of Agricultural and Resource Economics*, 52(4), 419-432.

[36] Taddese, G. (2001). Land degradation: a challenge to Ethiopia. *Environmental management*, 27(6), 815-824.

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

[38] United Nations, Department of Economic and Social Affairs, Population Division (2015). World Population Prospects: The 2015 Revision, Key Findings and Advance Tables. *Working Paper No. ESA/P/WP.241.*

[39] Weitzman, M.L., 1992. On diversity. *The Quarterly Journal of Economics*, 107(2), 363-405.

[40] World Bank. 2015. *World Development Indicators 2015*

[41] World Bank (2007a). Ethiopia Accelerating Equitable Growth Country Economic Memorandum: Part I: Overview. Report No. 38662-ET.

[42] World Bank (2007b). Ethiopia Accelerating Equitable Growth Country Economic Memorandum: Part II: Thematic Chapters. Report No. 38662-ET

Chapter 4

SWC Technology adoption and labour allocation: Implications for impact evaluation and policy. A case study of Ethiopia

Abstract

Soil and Water conservation (SWC) technologies have long been viewed as part of a solution to increase the resilience of agriculture to climate change. Recent research has also shown that they can also have a positive impact on agricultural yields. In this paper, I focus on the labour impacts of adopting SWC technologies (mostly bunds and artificial waterways) in Ethiopia. I use an endogenous Switching Regression Model (ESRM) and find that such technologies lead to large increases in plot-level adult and child labour, with estimated impacts ranging from 25% to 36%.

These findings are important for a number of reasons. First, it is important for policy-makers to understand the potential labour impacts of the widespread adoption of SWC technologies as these can affect other policy-priority areas (e.g. education). Second, large and significant labour impacts question the appropriateness of using econometric methods that assume no changes in inputs as a consequence of adoption (e.g. PSM or ESRM). Finally, these labour impact estimates also provide a plausible explanation to the puzzling result in Di Falco et al. (2011). Using the same dataset, the authors find that non-adopters have higher predicted gains from adoption. I find that non-adopters already work substantially more days than adopters and, as a result, may be less willing to adopt a labour-intensive technology.

JEL classification: J22, Q10, Q12, Q16, Q56

4.1 Introduction

Achieving economic development in Africa cannot be disentangled from the performance of the agricultural sector (Collier and Dercon, 2014; Diao et al., 2010). Consequently, achieving protracted productivity growth in the agricultural sector in an environmentally sustainable manner is key. As such, development actors have promoted agricultural technologies believed to hold the potential to increase agricultural production in a more environmentally sustainable way.

Soil and Water Conservation (SWC) technologies constitute an example of a set of technologies/practices which could ostensibly yield such a win-win situation. Consequently, in an attempt to gauge their success, a substantial amount of literature has been devoted to analysing the impacts of adopting such technologies. So far, most impact evaluation studies have focused on production and productivity metrics (Di Falco, 2011; Kato et al., 2011; Gorst et al., 2015). However, the productivity impact estimates typically rely on methods that implicitly assume that adopting the technology does not change the inputs used. In this paper, I focus on the potential impacts of SWC adoption on labour, where this assumption is unlikely to hold.

The three most adopted SWC structures in our dataset (soil bunds, stone bunds and artificial waterways[1]) can theoretically increase productivity. Bunds can potentially reduce erosion through reducing the velocity of run-on and waterways can drain excess water (Haile et al., 2006). However, the installation and maintenance of bunds and waterways has been found to be labour-intensive (Haile et al., 2006; Shiferaw and Holden, 1999; Pretty, 1999; Bekele and Drake, 2003). As a result, I empirically estimate the impact of the adoption of SWC technologies on the levels of both child and adult family labour. Understanding the existence and potential magnitude of the labour impact is likely to be important for a number of reasons.

First, from a methodological perspective, most studies estimating the productive impact of the adoption of SWC technologies use methods that rely on the assumption that the inputs do not change as result of adoption of a certain technology. As a result, if it is found that significant changes in inputs occur following the adoption of a technology, this may call into

[1]These three structures account for more than 85% of the installed structures in our dataset.

question the accuracy of impact estimates obtained.

Second, it is important for policy-makers to understand the full range of impacts of the widespread adoption of SWC technologies. For instance, an impact on the levels of child labour used may ultimately affect educational outcomes. As a result, it is important for policy-makers to understand the potential cross-sectoral implications of the widespread adoption of agricultural technologies, so as to devise policies to mitigate them.

Finally, the predicted impacts of adoption on input use may be a key determinant/deterrent of adoption of an agricultural technology. Better understanding the impacts of technology adoption on non-productivity metrics may also provide a plausible channel to explain some puzzling findings in the literature. Specifically, previous research (Di Falco et al., 2011) using the same dataset has found that non-adopters of a technology have highest predicted gains from adoption. This finding is puzzling, runs against economic rationality and is largely left unexplained in Di Falco et al. (2011). In this paper, I argue that estimated impacts on inputs may provide a plausible explanation for this finding and use labour as an illustration.

In order to theoretically motivate my question, I borrow and extend very slightly the model proposed by Fernandez-Cornejo et al. (2005). This model, based on the Agricultural Household model, captures the idea that the decision to adopt a new technology depends both on the expected benefits (gains in productivity) as well as on a range of other factors, such as monetary (additional inputs) and non-monetary (e.g. increased labour requirements) costs as well as household preferences. The model highlights the importance of studying the impacts on non-productivity metrics. Additionally, the model conceptually highlights how, despite higher predicted productivity gains, households may not adopt.

Empirically, I use a plot-level dataset of households in the Nile Basin of Ethiopia (Ringler and Sun 2010) and I use an Endogenous Switching Regression Model to estimate the impacts of adoption on the levels of adult and child labour. Robustness checks are performed using both an IV regression and an IV probit regression.

Overall, I find evidence that soil and water conservation technologies are associated with significant and large (31.4%) increases in adult labour. This result is robust across all the different specifications used and estimated impacts range from 25% to 36%, depending on the specification. In terms of child labour, I find no evidence of an increase in probability of using child labour. However, we find evidence of a large and significant impact (about 29%) in the levels of child labour. The magnitude of the coefficients is similar in the IV setting with estimates ranging from 30% to 34%, although the results are no longer statistically significant.

The results also reveal different patterns of labour impacts between adopters and non-adopters. Non-adopters display higher impacts on child labour while adopters display a higher impact on adult labour. This result can be explained by the fact that adults in non-adopting households already work longer hours. As a result, children are likely to have to bear a higher share of the additional labour requirements induced by the adoption of a labour-intensive technology. These findings also provide a plausible channel to explain why, despite higher predicted gains in productivity, non-adopters do not adopt the technology, as found in Di Falco et al. (2011). Given that adults in households that have plots where the technology was not adopted already work more days, they may be less willing to engage in a labour-intensive technology.

The rest of the paper is structured as follows. Section 2 reviews the literature on the Agriculture-Development nexus and the impact evaluation of SWC technologies. Section 3 discusses the theoretical model used to derive some of our theoretical predictions. Section 4 discusses the empirical methodologies used. Section 5 discusses the data. Section 6 presents the results. Finally, section 7 concludes.

4.2 Agriculture and Economic Development

The importance of agriculture in the economic development process is beyond dispute and it has been repeatedly linked to a number of socio-economic outcomes ranging from the incomes of a large portion of the world's population (Schultz, 1980), to various health outcomes (Hawkes and Ruel, 2006). The many ramifications of agricultural performance implied that it has occupied a central role in economic development theory, at least since Lewis 1954, when agriculture was viewed as an indispensable support for the engine of growth (industry) of a country (Lele and Mellor, 1981).

Since, economic thought appears to be adjusting to a new paradigm and the reigning perception among the various actors of development is that Agriculture constitutes something akin to a magic bullet, it being directly linked to poverty, food security, gender equality, education and environmental sustainability (Byerlee et al., 2009). This is epitomized by the preponderance of rural poverty, recurring national and international food crises (culminating in the most notable example in 2008), its importance as the main source of female employment and its environmental importance (responsible for 85%, 40% and 30% of the developing world's freshwater use, land use and greenhouse gas emissions, respectively). In addition to this, studies have repeatedly shown that child employment in agriculture is a strong negative determinant of both participation in education and educational achievement (Beegle et al., 2009; Heady, 2000).

However, few would argue that the future is without challenges. First, projected future demographic changes (close to 10 billion people by 2050; UN, 2015) pose a significant threat to agriculture, through its effects on food demand. As argued by Godfray et al. (2010), this increase in food demand is likely to increase the competition for natural resources, which in turn, will affect our ability to produce food. Secondly, the impacts of climate change are largely predicted to be negative (Mendelson, 2008). This is particularly important in the case of Africa, which is expected to witness the highest population growth (UN, 2015) and where crop-specific losses are predicted to be between -8% and -22% (Schlenker and Lobbel, 2010).

Together, these factors suggest that the model that has been responsible for sustained growth in agricultural output since the middle of the 20th century may not be sustainable. This has led to calls for a radical change in how food is produced (Godfray et al., 2010). Specifically, a more sustainable model of agriculture, capable of meeting global food needs in a more environmentally sustainable way, is an imperative condition to a more prosperous future.

Given this context, the promotion of more environmentally responsible agricultural practices, such as the promotion of SWCT (Soil and Water Conservation Technologies), appear a step in the right direction. In theory, adoption of SWCT may lead to lower levels of environmental degradation, while maintaining or even increasing agricultural productivity.

4.2.1 SWC technologies in Ethiopia

Description of the main SWC technologies in the sample

Soil degradation is widely considered an important issue in Ethiopia and erosion (especially wind and water erosion) is seen as the major cause of soil degradation. While both types of erosion (wind and water) occur as a natural geological process (Haile et al., 2006), anthropogenic factors are believed to accelerate the rate of erosion. Chief among these factors are the over-utilization of land for agriculture, overgrazing and deforestation (Haile et al., 2006), all of which tend to lead to reduced vegetative cover, thus increasing the risk of erosion.

As explained in Haile et al. (2006), soil erosion is likely to have productive implications since both water and wind erosion tend to lead to a less favourable distribution of soil. Typically, topsoil (richer in minerals and organic matter) is removed. As a result, even a minor loss in topsoil can substantially reduce the productivity of the soil. Results from experimental studies in Canada highlight the importance of topsoil. Specifically, Larney et al. (2000) found that removing 20 cm of topsoil reduced productivity by as much as 53%. In the Ethiopian Highlands, Tadesse (2001) estimates annual losses of topsoil to amount to 1.5 billion tons, costing Ethiopia an estimated 1-1.5 million tons of grain annually.

However, large losses due to erosion can be mitigated, at least partly, by using certain farming practices and/or installing certain structures capable of reducing the rate of erosion. This is not new in the Ethiopian context and, at least since 1973/74, there have been efforts to promote such structures (Osman and Sauerborn, 2001). Soil erosion was largely perceived to have compounded the effects of the 1973/74 drought that led to a famine (Osman and Sauerborn, 2001). This prompted the Ethiopian Government to promote the adoption of SWC practices and the construction of structures to mitigate the negative impacts of erosion. In the early stages, these efforts led to large increases in the area covered by conservation practices from an initial 7,000 ha in 1973 to about 163,000 ha in 1983 (Osman and Sauerborn, 2001).

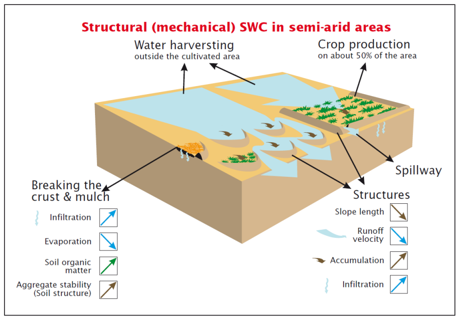
A large number of agronomic practices and physical structures fall under the definition of SWC technologies/practices and these differ widely from one another in the inputs required, their productive potential and their applicability to specific settings. However, Haile et al.

(2006) argue that SWC practices can be broadly subdivided into three broad types. These include agronomic measures (e.g. tillage, contour planting), vegetative measures (e.g. grass strips) and structural measures (e.g. bunds and waterways). In our sample, the majority (over 85%) of plots have adopted structural measures (mostly bunds and waterways) and very few adopt vegetative measures such as grass strips and trees. As a result, we will focus mostly on bunds and waterways and will only briefly describe one vegetative cover measure (Grass strips).

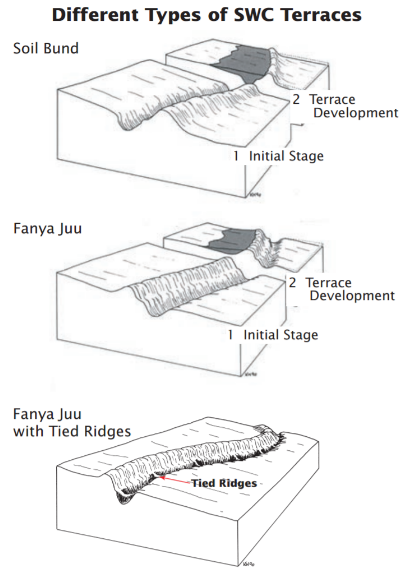
The main purpose of soil and stone bunds (illustrated in figure 4.1) is to retain water on the cropping area (runon) while draining excess water as necessary. Panel (a) of figure 4.1 illustrates how these structures reduce the velocity of the runoff and reduce slope length, allowing for greater infiltration (Haile et al. 2006). As a result, bunds can reduce the risk of erosion and increase soil quality and soil moisture (Haile et al., 2006; Kato et al., 2011). However, different types of bunds exist and a distinction needs to be drawn between soil and stone bunds. Stone bunds are more permeable, thus allowing for better drainage of excess water (Haile et al., 2006). Conversely, soil bunds are able to retain more water, but tend to be more prone to both wind and water erosion and require more regular maintenance (Haile et al., 2006). As a result, soil bunds tend to be preferred in areas with low moisture (e.g. arid areas) and/or in areas with limited supply of stones. Conversely, stone bunds are likely to be preferred in areas where excess water is likely to be a concern (sub-humid and humid areas). Overall, Haile et al. (2006) highlights that bunds (especially soil bunds) tend to characterized by low durability and large maintenance requirements. Kato et al. (2011) add that they tend to be costly and hard to build[2].

[2] According to Kato et al. (2011), in our sample, when these structures are too costly and/or materials are unavailable, vegetative measures are adopted instead.

Figure 4.1: Graphical illustration of bunds



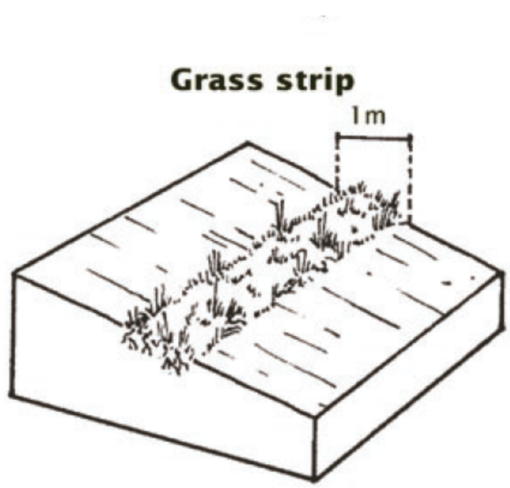
(a)SWC system in a semi-arid area
Source: Haile et al. (2006)



(b)Bunds and terraces
Source: Haile et al. (2006)

Vegetative measures increase soil cover, which helps to control erosion, as illustrated in figure 4.2. According to Haile et al. (2006), vegetative measures provide fairly good erosion control (though not as good as bunds) but require low amounts of additional labour[3]. However, grass strips are prone to weed infestation and the harbouring of rodents (Haile et al., 2006).

Figure 4.2: Graphical illustration of Grass strips

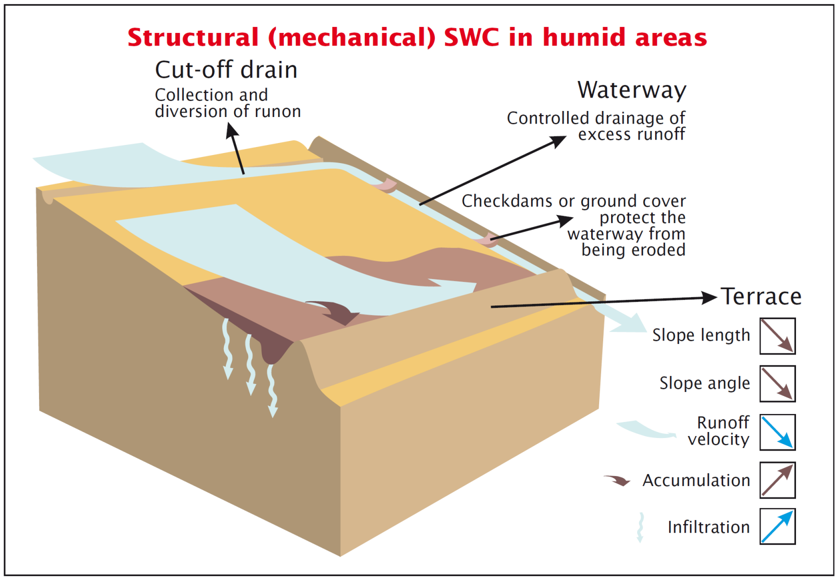


Source: Haile et al. (2006)

[3] According to Kato et al. (2011), the risk of erosion is reduced as grass strips reduce runoff velocity, which allows for better infiltration.

Regarding artificial waterways, their main purpose differs from that of either grass strips or bunds. As illustrated in figure 4.3, the main aim of artificial waterways is to safely conduct runoff (excess) water safely along specified pathways in the fields (Kato et al., 2011), often downhill into a river or stream (Haile et al., 2006).

Figure 4.3: Graphical illustration of artificial waterways



Source: Haile et al. (2006)

Given their different purposes, bunds and waterways are likely to suit different needs. Bunds (soil or stone) are likely to be preferred in areas where moisture deficiency is an issue, whereas waterways are likely to be preferred in areas subject to excess or extreme rainfall. This pattern is present in our dataset. Kato et al. (2011) show that bunds are adopted in 37% of plots in low rainfall areas (60% in Tigray) compared to 31% in high rainfall areas. The reverse pattern holds for waterways, where waterways are adopted in 41% of the plots in high-rainfall areas against 15% in low-rainfall areas.

Regarding labour requirements, a number of authors have highlighted that both waterways and bunds tend to be very labour intensive. For instance, Shiferaw and Holden (1999) use a modelling assumption of 100 person-days per hectare to install the soil conservation structures and an additional 20 person-days per year for maintenance. Bekele and Drake (2003) recommend between 70 to 150 person-days per hectare for different types of soil bunds in the Eastern Highlands of Ethiopia. In addition to this, a number of other studies also highlight the high labour requirements during installation and maintenance for a number of Soil and Water Conservation technologies (McCarthy, 2011; Bekele and Drake, 2003; Bewket, 2007; Pretty, 1999; Tesfaye et al., 2016), including bunds and waterways[4], though they do not provide estimates of the number of days. Finally, as mentioned by Haile et al. (2006), vegetative measures are likely to require lower amounts of labour compared to bunds and waterways.

Impacts of SWC technology adoption in Ethiopia: a brief review of the literature

To date, there have been a substantial number of studies assessing the impacts of SWC technologies in Ethiopia. However, most studies in Agricultural Economics have focused on the productive implications of SWC technologies, leaving aside potential impacts on input-use.

Kato et al. (2011) measure the impact of adoption of SWC technologies on mean yields and variance. The authors find that soil bunds have a positive yield effects in low rainfall areas, whereas planting trees and waterways have positive yield effects in high rainfall areas. Regarding the impacts on yield variance, the authors find that only soil bunds have a significant negative impact on the mean yield variance in low rainfall areas. A similar pattern is found in Kassie et al. (2008), who find that the adoption of soil bunds leads to a large increase in yield (about 25% of the mean value) in low rainfall areas, specifically in the Tigray region. The results by Kato et al. (2011) and Kassie et al. (2008) highlight that impacts are likely to differ depending on the amount of rainfall available in the area. However, both studies suffer from two weaknesses. Firstly, both assume a unique production function (one equation). This implies that, implicitly, the authors are assuming that both adopters and non-adopters have the same elasticities of inputs. This is questionable, since it is highly likely that the retention of moisture may enable better use of complementary inputs such as fertilizer. Secondly, as with the majority of other studies reviewed in this section, these results fail to shed light on the effect of the adoption of such structures on input-use, which is likely to be a key factor behind the decision to adopt of SWC technologies.

Deressa et al. (2009) investigate the determinants of adoption of SWCT at the household

[4]McCarthy (2011) mentions that soil and water conservation technologies entail large upfront costs, and in the cases of stone bunds, for example, both the initial construction and annual maintenance entails heavy labour requirements. In Bewket (2007), who looks at Ethiopia, 92% of the farmers surveyed argued that the introduced SWC structures (which included bunds, waterways and fanya juu) required too much labour to implement. Pretty (1999) argues that artificial waterways and dams are also labour-intensive

level using the same dataset and a multinomial logit specification. The authors find that the level of education, the gender and the age of the head of the household all positively affect the probability of adoption of some technologies. In addition to this, non-farm income, farm income as well as the agroecology were significant determinants of adoption. An interesting result in this paper is the effect of climate change information, whose sign depended on the technology[5].

Di Falco et al. (2011) use the same plot-level dataset and estimate the impact of adopting SWCT on yields using an Endogenous Switching Regression model. The authors use information sources as instruments[6] and estimate large positive effects of adoption on yield, averaging 181 kg/ha for adopters and almost 300 kg/ha for non-adopters. The authors find that non-adopters stand to gain more from adoption. However, the authors do not explicitly propose a channel to explain this finding, though they highlight the importance of increasing climate awareness and access to credit.

Two studies that have looked at the effects on input-use are Zeng et al. (2013) and Kassie et al. (2015). Zeng et al. (2013) look at the impacts of adopting improved varieties on yields and costs in Ethiopia. The authors find that improved varieties were associated with both higher yields and higher costs. As a result, poorer farmers tended to benefit least from the adoption of these varieties, given their small land sizes. Kassie et al. (2015) who use data from Malawi, show that a number of technologies[7] are often associated with statistically significant changes in input application (fertilizer and pesticide application in their case), thereby implying an increase in costs.

As such, so far there appears to be positive returns in terms of agricultural production. However, the fact that the gains for non-adopters are so high, often higher than non-adopters is puzzling. More so, since this finding is not limited to Ethiopia and/or the adoption of SWC technologies. Using data from Pakistan Gorst et al. (2015) find that non-adopters consistently have higher gains from adoption than adopters over all the crops that the authors analyse. Khonje et al. (2014) show that in Zambia, non-adopters would have the highest impact

[5]For some technologies climate information was associated with a higher probability of adoption whereas, in other cases, it was associated with a lower probability of adoption.

[6]Including farmer-to-farmer extension, government extension, climate information, neighbourhood information and radio information.

[7]The authors focus on IAPs (Improved agronomic practices).

of adoption of improved maize varieties on food security. However, while this pattern is not uncommon, few studies have attempted to answer this and/or have analysed channels through which this may happen.

One of the first studies that can possibly provide a plausible explanation is that of Zeng et al. (2013). The authors find that, while improved varieties were associated with higher levels of yields, they were also associated with an increase in costs. As a result, poorer farmers tended to benefit least from the adoption of these varieties, given their small land sizes. Another potential explanation could be the potential changes in input choice found in Kassie et al. (2015), which shows that a number of technologies are often associated with a positive change in inputs, thereby implying an increase in costs.

However, so far, to my knowledge, the labour channel has not been analyzed. Agriculture is generally considered to be an arduous activity and some of these technologies tend to be quite labour-intensive, especially at the early stages of adoption. By way of example, Shiferaw and Holden (1999) use a modelling assumption of 100 person-days per hectare to install the soil conservation structures and an additional 20 per year for maintenance. Bekele and Drake (2003) recommend between 70 to 150 person-days per hectare for different types of soil bunds in the Eastern Highlands of Ethiopia. However, a number of other studies also highlight the high labour requirements during installation and maintenance for a number of Soil and Water Conservation technologies (McCarthy, 2011; Bekele and Drake, 2003; Bewket, 2007; Pretty, 1999; Tesfaye et al., 2016).

As a result, given the likely installation and maintenance labour requirements, it seems perhaps far-fetched to assume that there will be no labour re-adjustment emanating from the adoption of such technologies. As such, farmers, either through their knowledge of their land or through observing other farmers, tend to be aware (or at least have an idea) of the likely impacts of adopting a new technology on labour requirements. These labour requirements can then have effects on aspects such as adult labour, child labour and time availability for off-farm activities. As such, it seems possible that farmers may self-select according to the anticipated (child and adult) labour requirements. In fact, this is perhaps echoed in a qualitative study by Bewket (2007) who survey farmers in a watershed in the north-western highlands of Ethiopia. Labour shortages and unavailability of labour were viewed as key aspects discouraging the adoption of SWC technologies.

In this paper I argue that understanding and quantifying the labour impact, however, is not trivial. If there is a positive effect on total farm labour, this may shift the decision to employ children and/or increase the time children spend in agricultural activities. As such, this is a policy-relevant question, since the amount of child labour may have cross-sectoral implications in areas such as education (Abafita and Kim, 2014). Moreover, Colmer (2013) finds that climate change affects the hours spent by children on farming and labour activities. The author argues that the increase in child labour is a way of mitigating future anticipated climate shocks. However, no channel is proposed regarding the mechanism through which this occurs. We argue that technology adoption may be a plausible channel.

The next section introduces an agricultural household model which tries to disentangle the relationship between child labour, off-farm work, farm-work and technology adoption in the same framework. It also serves to motivate further the use of empirical methods capable of incorporating self-selection based on unobservables.

4.3 Theoretical Model

In order to further motivate the research question, we will look at the adoption decision using a slightly modified version of the model presented by Fernandez-Cornejo et al. (2005), which integrates the adoption decision into the Agricultural Household model. The main difference is that we consider a two-member household where the household has to make a choice regarding the labour and education trade-off of children. Specifically, the model highlights two features. First, it highlights the importance of non-productivity metrics when considering the impacts of adoption of SWC technologies. Secondly, it highlights how the adoption process may depend on a number of factors not observed by the researcher, which motivates the use of our empirical methodology explained in Section 4. In this section, we present the main equations from the model. The full model, the set of assumptions and first-order conditions are available in Appendix A.

In this model a household is composed of two members (or groups of members), one adult and one child. The household seeks to maximize its utility subject to a number of constraints, namely an income constraint, a technology constraint and two individual time constraints. For simplicity, only one non-labour input is assumed and I restrict the availability of education to children and the availability of off-market opportunities to adults (i.e. adults can work off-farm but cannot go to school, whereas the opposite holds for children).

Algebraically, the constraints can be denoted as follows:

$$MaxU = U(Y_1, Y_2, L_1, L_2, E_2) \tag{4.1}$$

subject to:

$$p_y(Y_1 + Y_2) + p_e E_2 = p_q Q - w_x X + w_m M_1 \tag{4.2}$$

$$Q = Q[X(\tau), F_1(\tau), F_2(\tau), \tau], \;\; \tau \geq 0 \tag{4.3}$$

$$T_1 = F_1(\tau) + M_1 + L_1, \;\; M_1 \geq 0 \tag{4.4}$$

$$T_2 = F_2(\tau) + E_2 + L_2, \;\; E_2 \geq 0 \tag{4.5}$$

Intuitively, equation 4.1 states that the household utility depends on the choices of labour, leisure, education and consumption of the household. The household is composed of two members, an adult (subscript 1) and a child (subscript 2). Specifically, I assume that the utility function depends positively on the amount of leisure and consumption of both members (L_1, L_2, Y_1 and Y_2). Furthermore, the utility function also depends on the amount of time the child spends on education (E_2).

Equation 4.2represents the income constraint and states that the total household expenditure on consumption goods ($p_y(Y_1 + Y_2)$[8]) and education ($p_e * E_2$[9]) cannot exceed the total

[8] where p_y denotes the market price of good Y. Y_1 and Y_2 denote the quantities of good Y consumed by the adult and the child, respectively.

[9] where p_e denotes the market price of education and E_2 denotes the time spent in education by the child in the household.

household income, which is composed of farm profits ($p_q * Q - w_x * X^{10}$) and off-farm income ($w_m * M_1{}^{11}$).

Equation 4.3 is the main novelty proposed by Fernandez-Cornejo et al. (2005).It represents the technological constraint (a concave, continuous, twice differentiable production function). Key to the technological constraint is the parameter τ, which can be interpreted as a measure of intensity of adoption of SWC technologies. Specifically, $\tau = 0$ represents a non-adopter and $\tau > 0$ suggests that a SWC technology has been adopted, with higher values of indicating a more intensive adoption of these technologies. In our model, the choice of technology affects production through two channels. First, there is pure productivity shift that occurs irrespective of the inputs. Secondly, the technology also affects input-use since different technologies will have different elasticities of production.

The final two constraints, equations 4.4 and 4.5 represent the time constraints for both members of the household. Adults allocate their total time endowment between leisure (L_1), off-farm work (M_1) and farm work (F_1) activities. The total time endowment of children is allocated between leisure (L_2), education (E_2) and farm-work (F_2).

Substituting equation 4.3 into equation 4.2, a technology-constrained version of the cash constraint of the outcome can be obtained and is given by the following equation:

$$p_y(Y_1 + Y_2) + p_e E_2 = p_q Q[X(\tau), F_1(\tau), F_2(\tau), \tau] - w_x X(\tau) + w_m M_1 \tag{4.6}$$

Given these equations, the Lagrangean is given by the following equation:

$$\begin{aligned}\mathcal{L} = U(Y_1, Y_2, L_1, L_2, M_1, E_2) + \lambda[p_q Q[X(\tau), F_1(\tau), F_2(\tau), \tau] - w_x X(\tau) + w_m M_1 \\ -p_y(Y_1 + Y_2) - p_e E_2] + \mu[T_1 - L_1 - F_1(\tau) - M_1] + \phi[T_2 - L_2 - F_2(\tau) - E_2]\end{aligned} \tag{4.7}$$

The full set of first-order conditions and assumptions is available in Appendix A. In this

[10] p_q and w_x represent the market prices of agricultural outputs and non-labour agricultural inputs, respectively. Q and X represent the total agricultural production and the total quantity of non-labour inputs used in the production process.

[11] w_m represents the wage rate for off-farm labour and M_1 represents the total time spent on off-farm labour.

section, I assume that the model is separable and focus on the first-order condition related to the adoption decision, summarized in equation 4.8, below:

$$\begin{aligned}\frac{\partial \mathcal{L}}{\partial \tau} = \lambda \Big[p_q \Big[\frac{\partial Q}{\partial X}\frac{\partial X}{\partial \tau} + \frac{\partial Q}{\partial F_1}\frac{\partial F_1}{\partial \tau} + \frac{\partial Q}{\partial F_2}\frac{\partial F_2}{\partial \tau} + \frac{\partial Q}{\partial \tau} \Big] - w_x \frac{\partial X}{\partial \tau} \Big] \\ - \mu \frac{\partial F_1}{\partial \tau} - \phi \frac{\partial F_2}{\partial \tau} \leq 0; \;\; \tau \geq 0; \;\; \tau \Big(\frac{\partial \mathcal{L}}{\partial \tau} \Big) = 0\end{aligned} \tag{4.8}$$

Using equation 4.8, the following adoption decision can be derived:

$$p_q \frac{dQ}{d\tau} = w_x (\frac{\partial X}{\partial \tau}) + \frac{\mu}{\lambda} (\frac{\partial F_1}{\partial \tau}) + \frac{\phi}{\lambda} (\frac{\partial F_2}{\partial \tau}) \tag{4.9}$$

The left-hand side of equation 4.9 represents the gains from adoption (increased revenues from productivity gains) while the right-hand side represents the costs (or savings, if partial derivatives are negative) associated with the adoption of the technology. More importantly, equation 4.9 highlights three important aspects.

First, equation 4.9 highlights clearly the main hypothesis of this paper. We are essentially testing whether $\frac{\partial F_1}{\partial \tau} = 0$ and $\frac{\partial F_2}{\partial \tau} = 0$. This is important because most impact evaluation methodologies used in cross-sectional data assume that the adoption decision has no impact on input use (i.e. they assume that $\frac{\partial X}{\partial \tau} = 0$, $\frac{\partial F_1}{\partial \tau} = 0$ and $\frac{\partial F_2}{\partial \tau} = 0$).

Second, the equation highlights the importance of studying non-productivity metrics of technology adoption. Specifically, it provides a plausible channel to explain why households may not adopt a given technology, even when associated gains are large. As shown in equation 4.9, even if a given technology is associated with large production gains ($\frac{dQ}{d\tau} > 0$ and large), it will only be adopted if the gains compensate the changes in costs, both monetary (increase in inputs) and non-monetary (time spent on-farm) [i.e. if $p_q(\frac{dQ}{d\tau}) > w_x(\frac{\partial X}{\partial \tau}) + \frac{\mu}{\lambda}(\frac{\partial F_1}{\partial \tau}) + \frac{\phi}{\lambda}(\frac{\partial F_2}{\partial \tau})$].

Finally, equation 4.9 also motivates our preference for empirical methods taking into account unobservable heterogeneity. Specifically, technology adoption is likely to depend on a number of factors such as households expectations regarding input use ($\frac{\partial X}{\partial \tau}$, $\frac{\partial F_1}{\partial \tau}$ and $\frac{\partial F_2}{\partial \tau}$) as well as a household's valuation of their time ($\frac{\mu}{\lambda}$and$\frac{\phi}{\lambda}$). These factors are known (or predicted) to the household adopting the technology. However, the researcher does not observe them. This drives the choice the methodology, explained in the next section.

4.4 Empirical Methodology

4.4.1 The impact evaluation problem and selection biases

In impact evaluation, the main objective is to understand what is the effect (or impact) of a certain treatment (D) on a certain outcome of interest (Y). The main challenge to estimating the impact is that researchers are unable to simultaneously observe the outcome for individual i under adoption (Y_{i1}) and non-adoption (Y_{i0}). Formally, we can represent this by the equation below:

$$Y_i = Y_{i1} * D_i + Y_{i0} * (1 - D_i) \tag{4.10}$$

Consequently, we need to find a suitable counterfactual. Specifically, we have to look for the best possible estimate of the outcome of interest for the status for which individual i is not observed. Supposing that such counterfactual exists and is correctly estimated, different measures of impact can then be obtained. The most common measures are 1) the ATT (Average Treatment on the Treated), which measures the average treatment effect for those who have adopted, 2) the ATU (Average Treatment on the Untreated), which measures the Average Treatment Effect for the untreated, and 3) the ATE (Average Treatment Effect) which measures the average treatment effect for the full sample. The latter measure is a weighted average of the ATT and the ATU. Formally, these can be written as:

$$ATE = E[Y_{i1} - Y_{i0}] \tag{4.11}$$

$$ATT = E[Y_{i1} - Y_{i0}|D_i = 1] \tag{4.12}$$

$$ATU = E[Y_{i1} - Y_{i0}|D_i = 0] \tag{4.13}$$

The issue in non-experimental studies is related to self-selection. In other words, individuals may be observed in the treatment status that is most favourable for them. This implies that the outcome is no longer independent of the treatment status, as individuals choose the treatment status based on their expected outcome, as shown in the equation below:

$$y_0, y_1 \not\perp\!\!\!\perp D \tag{4.14}$$

If randomization is used and does not fail, the treatment status is independent of the outcome because, probabilistically speaking, individuals in both groups are balanced in both their observables and unobservable covariates.

In non-experimental settings, however, practitioners have to revert to alternative methods that, under certain assumptions, circumvent the selection bias and provide reliable estimates. Two key assumptions that are widely used in methods such as Propensity Score Matching (PSM), are that of conditional independence and common support. The basic idea behind propensity score matching is that, based on observable variables alone, the matching process makes the selection into treatment as good as random, as shown in the equations below:

$$(y_1, y_0)|X \perp\!\!\!\perp D \tag{4.15}$$

$$0 < Pr(D = 1|X) < 1 \tag{4.16}$$

However, as argued by Heckman and Vytlacil (2007), this is a very strong assumption. It requires that the researcher possesses and uses all the relevant information available and used by the individual when selecting into the treatment. As highlighted in the theoretical model, we do not believe this holds in this case[12]. As a result, I opt for a model that allows for unobserved heterogeneity, namely an Endogenous Switching Regression Model (ESRM).

[12] I argue that aspects such as time valuation and predicted changes in inputs are likely to be determinant factors. These are aspects I do not observe.

4.4.2 Endogenous Switching Regression Model

The Endogenous Switching Regression model consists of two stages. In the first stage, a farmer will adopt a technology based on his expected benefits (expected outcome) from adoption. These benefits depend on both observable characteristics (W_i) and unobservable characteristics (u_i^d). Formally:

$$D_i^* = \beta W_i + u_i^d \tag{4.17}$$

$$D_i^* = \begin{cases} 1 & \text{if } D_i^* > 0 \\ 0 & \text{otherwise} \end{cases} \tag{4.18}$$

As a result, a farmer will only adopt a certain technology if there are net benefits of adoption. Based on the adoption decision, a different equation is estimated for each state:

$$Y_i = \begin{cases} Y_{i1} = \beta_1 X_i + \epsilon_{i1} & \text{if } D == 1 \\ Y_{i0} = \beta_0 X_i + \epsilon_{i0} & \text{if } D == 0 \end{cases} \tag{4.19}$$

This equation summarizes one of the main strengths of this empirical approach. By estimating two separate equations, the method allows for the outcome variable (in our case labour) to react differently to the explanatory variables, depending on the adoption status. In the case of a labour-intensive technology (such as bunds), we would expect the coefficient on plot size to be higher when the technology is adopted. In other words, we would expect households to spend more days per hectare cultivating a plot where bunds have been adopted. In order to help with identification, we also include one variable in the adoption equation that omitted from the labour equation. We will discuss our choice of instrument in the next section.

Another feature of the Endogenous Switching Regression model is that it assumes a non-zero correlation between the error term of the adoption equation and the error terms of the outcome equations. This is shown in the matrix below, which provides the variance-covariance matrix for the endogenous switching regression model:

$$\Sigma = \begin{bmatrix} \sigma_u^2 & \sigma_{u\epsilon 1} & \sigma_{u\epsilon 0} \\ \sigma_{u\epsilon 1} & \sigma_1^2 & 0 \\ \sigma_{u\epsilon 0} & 0 & \sigma_0^2 \end{bmatrix}$$

Where σ_u^2 represents the variance of the error term in the selection equation, σ_1^2 and σ_0^2 represent the variance of the outcome equations under adoption and non-adoption, respectively. The terms $\sigma_{u\epsilon 1}$ and $\sigma_{u\epsilon 0}$ represent the covariance between the selection equation and outcome regimes under adoption and non-adoption, respectively. The errors follow a trivariate normal distribution with mean 0 and the variance-covariance matrix Σ.

In practice, this implies that the outcome equations need to be adjusted for the sample selection in the following way:

$$E[\epsilon_{i1}|D = 1] = \sigma_{w1}\frac{\phi(\beta W_i)}{\Phi(\beta W_i)} \tag{4.20}$$

$$E[\epsilon_{i0}|D = 0] = \sigma_{w0}\frac{\phi(\beta W_i)}{1 - \Phi(\beta W_i)} \tag{4.21}$$

Where the terms $\frac{\phi(\beta W_i)}{\Phi(\beta W_i)}$ and $\frac{\phi(\beta W_i)}{1 - \Phi(\beta W_i)}$ are also known as the inverse Mills ratios. In order to test for the presence of selection bias, two coefficients (ρ_1 and ρ_0) are calculated. If they are statistically significant, this suggests the presence of selection bias. These coefficients are calculated using the following expression:

$$\rho_1 = \frac{\sigma_{u1}^2}{\sigma_u \sigma_1} \tag{4.22}$$

$$\rho_0 = \frac{\sigma_{u0}^2}{\sigma_u \sigma_0} \tag{4.23}$$

Given the equations above, we can then define the conditional outcomes for adopters and non-adopters as follows:

$$E(Y_{i1}|D = 1) = \beta_1 X_{i1} + \sigma_{u1}\frac{\phi(\beta W_i)}{\Phi(\beta W_i)} \tag{4.24}$$

$$E(Y_{i0}|D=0) = \beta_0 X_{i0} + \sigma_{u0}\frac{\phi(\beta W_i)}{1-\Phi(\beta W_i)} \tag{4.25}$$

using this framework, the counterfactual outcomes are given by:

$$E(Y_{i1}|D=0) = \beta_1 X_{i1} + \sigma_{u1}\frac{\phi(\beta W_i)}{1-\Phi(\beta W_i)} \tag{4.26}$$

$$E(Y_{i0}|D=1) = \beta_0 X_{i0} + \sigma_{u0}\frac{\phi(\beta W_i)}{\Phi(\beta W_i)} \tag{4.27}$$

And then we obtain the ATT by simply taking the difference between equations 4.24 and 4.27 and the ATU is obtained by taking the difference between equations 4.26 and 4.25. The corresponding equations are:

$$ATT = E(Y_{i1}|D=1) - E(Y_{i0}|D=1) = X(\beta_1 - \beta_0) + (\sigma_{u1} - \sigma_{u0})\frac{\phi(\beta W_i)}{\Phi(\beta W_i)} \tag{4.28}$$

$$ATU = E(Y_{i1}|D=0) - E(Y_{i0}|D=0) = X(\beta_1 - \beta_0) + (\sigma_{u1} - \sigma_{u0})\frac{\phi(\beta W_i)}{1-\Phi(\beta W_i)} \tag{4.29}$$

4.5 Data and Choice of Instrument

4.5.1 Data

In this paper, I use the Ethiopia Nile Basin Climate Change dataset (Ringler and Sun 2010). The dataset is publicly available from the International Food Policy Research Institute website[13]. The data relies on survey data collected from 1,000 farm households, over 20 woredas spread over different agro-ecological zones[14] in the Nile Basin in Ethiopia in 2005.

This dataset has three main strengths. First, data is collected at the plot level. Second, the dataset contains detailed information on the adoption of various soil and water conservation practices. Finally, a large number of plots are sampled (more than 3,000 plots). The main

[13] The survey was conducted by the Ethiopian Development Research Institute (EDRI), in collaboration with the International Food Policy Research Institute (IFPRI). Funding for the survey was provided by the Federal Ministry for Economic Cooperation and Development (Germany). The project forms part of the Consultative Group on International Agricultural Research (CGIAR)'s Challenge Program on Water and Food (CPWF).

[14] Dega, Woina Dega, Kolla and Berh

weakness of this dataset is that it is a cross-section. As a result, I am not able to control for factors such as household fixed-effects. However, throughout the paper, I will show that, for a sub-sample of partial adopters (i.e. those that adopt in some plots but not in others), for which we can use household fixed-effects, our results are robust to the inclusion of fixed-effects.

For the purposes of this study, I am mainly interested in SWC technologies. In particular, it is noteworthy that waterways and/or bunds were adopted in over 85% of the plots where SWC technologies were adopted. The analysis is limited to plots that cultivate legumes (Beans, Chickpeas, Cowpeas, Field Peas and Lentils) and/or cereals (Barley, Maize, Millet, Oats, Teff, Wheat, Sorghum and Finger Millet/Dagusa) and data for both the Meher and Belg season is used, though most of the data is from the Meher season. As a result, 3,633 are used in the analysis. The steps involved in the data preparation are described Appendix B.

In terms of labour, I focus on family labour which, in the African and Ethiopian accounts for the vast majority of on-farm labour. Previous research by Dillon and Barret (2017) shows that in the five analysed Living Standards Measurement Study (LSMS) surveys (covering five countries, namely Ethiopia, Malawi, Niger, Tanzania and Uganda) the proportion of work carried out by hired labour never exceeds 20% of total labour. A similar conclusion is reached by Palacio-Lopez et al. (2017). The results suggest that in all the analysed countries (Nigeria, Niger, Ethiopia, Malawi, Tanzania and Uganda) hired labour accounts for a small proportion of total labour. Specifically, in all countries at least 90% of the labour is provided by adult household members. In the Ethiopian context, Bachewe et al. (2016) find that in high-potential areas of Ethiopia, agricultural wage income accounted for about 7% of rural income and hired labour represented 7% of total labour[15]. These numbers are in line with the labour composition in our sample. In our sample, approximately 85% of the labour is provided by adult household members, 10% is provided by children[16] and only 5% is provided by hired labour. The average plot uses approximately 2 days of hired labour against 27 days of family labour (i.e. adult plus child labour).

Table 4.1 shows the summary statistics for the observations used in the sample. Table 4.1 shows that SWC technologies have been adopted in the majority of the sampled plots (2771 out of 3633). The summary statistics also suggest that adopters and non-adopters appear to

[15]The authors also find that only 1% of the farms relied solely on hired labour.
[16]Defined as a household member below 15 years old.

be very similar in terms of their demographic characteristics, both displaying an average of approximately three adults and three children per household. In terms of other household characteristics, adopters tend to be closer to the market and, with the exception of droughts, have had less exposure to natural disasters. TV ownership, however, tends to be higher among non-adopters. In terms of plot level characteristics, adopters differ from non-adopters. Adopters tend to have fewer but larger plots and display higher yields of both cereals and legumes. However, a larger proportion of plots where SWC technologies were not adopted are believed to be fertile.

Table 4.1: Descriptive Statistics by Adoption Status

	All			Non-Adopters			Adopters		
Variables	N	Mean	S.D.	N	Mean	S.D.	N	Mean	S.D.
Demographic Variables									
Number of Adults in the Household	3633	3.27	1.46	862	3.21	1.32	2771	3.29	1.50
Number of children in the Household	3633	3.30	1.74	862	3.31	1.89	2771	3.30	1.69
Household has a child under 6	3633	0.37	0.48	862	0.37	0.48	2771	0.36	0.48
Number of children under 6 in the Household	3633	1.03	0.96	862	1.03	0.97	2771	1.03	0.96
Household has a child between 6 and 11	3633	0.28	0.45	862	0.30	0.46	2771	0.27	0.45
Number of children between 6 and 11 in the Household	3633	1.22	0.99	862	1.27	1.13	2771	1.21	0.95
Household has a child between 11 and 15	3633	0.32	0.47	862	0.36	0.48	2771	0.30	0.46
Number of children between 11 and 15 in the Household	3633	1.05	0.89	862	1.00	0.92	2771	1.06	0.87
Household Characteristics									
Is head of household literate (1 Yes)	3633	0.50	0.50	862	0.48	0.50	2771	0.51	0.50
Distance to market where output is sold	3633	5.61	3.95	862	5.94	3.49	2771	5.51	4.08
Has experienced at least one hail storm since 1994	3633	0.21	0.40	862	0.15	0.35	2771	0.22	0.42
Household owns a TV	3633	0.43	0.49	862	0.49	0.50	2771	0.41	0.49
Has experienced at least one flood since 1994	3633	0.16	0.37	862	0.08	0.28	2771	0.18	0.39
Has experienced at least one drought since 1994	3633	0.40	0.49	862	0.25	0.44	2771	0.45	0.50
Plot-level Information (Non-Labour)									
Area of plot (ha)	3631	0.55	10.38	861	0.39	0.32	2770	0.60	11.88
Yield (all crops)	3525	988.17	896.39	831	915.87	901.24	2694	1010.48	893.87
Yield (kg/ha) cereals only	2910	1015.00	914.68	667	955.14	939.77	2243	1032.81	906.53
Yield (kg/ha) legumes only	615	875.64	843.07	164	756.13	703.83	451	919.10	885.03
Number of plots used in the analysis owned by the household	3633	4.85	2.03	862	4.97	2.02	2771	4.81	2.03
Plot is highly fertile	3633	0.27	0.44	862	0.32	0.47	2771	0.25	0.44
Plot is steep	3633	0.05	0.22	862	0.02	0.14	2771	0.06	0.24
Plot is flat	3633	0.59	0.49	862	0.68	0.47	2771	0.56	0.50
Plot has medium depth	3632	0.47	0.50	862	0.37	0.48	2770	0.51	0.50
Adopt Soil and Water Conservation Practices	3633	0.76	0.43	862	0.00	0.00	2771	1.00	0.00
Labour									
Total Child Labour in all activities	3633	3.14	6.54	862	1.83	4.36	2771	3.55	7.03
Total adult family labour (days per plot)	3633	24.23	18.92	862	24.99	19.28	2771	23.99	18.80
Total Family labour used by the Household (days per household)	3633	112.23	74.47	862	123.66	86.45	2771	108.67	69.96
Average family labour per adult in the household	3633	37.52	25.61	862	41.50	29.27	2771	36.29	24.23
Total child labour used in all plots (days)	3633	14.42	26.21	862	9.30	18.30	2771	16.01	28.03
Average days of child labour per child (days per household)	3447	4.80	8.90	812	3.34	7.27	2635	5.25	9.30
Total hired adult Labour days all activities	3633	1.95	9.50	862	3.60	17.26	2771	1.44	4.96
Share of adult labour	3633	0.87	0.14	862	0.89	0.14	2771	0.86	0.14
Share of family labour	3633	0.96	0.09	862	0.95	0.10	2771	0.96	0.08
Share of hired labour	3633	0.04	0.09	862	0.05	0.10	2771	0.04	0.08
Share of child labour	3633	0.09	0.13	862	0.06	0.10	2771	0.10	0.13
Regional Distribution of Plots									
Tigray	3633	0.15	0.35	862	0.05	0.21	2771	0.18	0.38
Amhara	3633	0.41	0.49	862	0.31	0.46	2771	0.44	0.50
Oromiya	3633	0.28	0.45	862	0.36	0.48	2771	0.26	0.44
Benishangul Gumuz	3633	0.10	0.30	862	0.07	0.26	2771	0.11	0.32
SNNP	3633	0.06	0.24	862	0.22	0.41	2771	0.01	0.10

N refers to the total number of observations, S. D. refers to the Standard deviation.

Adopters and non-adopters also differ substantially in terms of labour, both in terms of total days worked on the plots as well as in the distribution of labour between adults and children. Specifically, plots that have adopted SWC technologies tend to use more child labour, but use less adult family labour. Households who did not adopt SWC technologies in their plots also report working, on average, about 15 more days (approximately 14% more). With regards to child labour, plots where SWC technologies were adopted use about twice the amount of child labour and households that adopt report using approximately 6.5 additional days of child labour (about 77% more).

Another pattern highlighted in Table 4.1 is the geographical adoption patterns. In Tigray, for instance, SWC practices are used in the vast majority (over 85%) of the surveyed plots. The opposite pattern holds in the SNNP (Southern Nations, Nationalities and Peoples) region, where SWC technologies are used in very few plots. This pronounced geographical pattern suggests that regional differences in agricultural suitability and farming practices are an important factor underlying the adoption decision. As such, throughout the paper, I will attempt to incorporate these differences in the estimation procedure, namely through the inclusion of village fixed-effects.

4.5.2 Choice of Instrument

As mentioned in section 4.4, the identification of the equations in the ESRM relies on the choice of a variable included in the adoption equation and excluded from the outcome equation. Specifically, it requires an instrument that is both highly correlated with the adoption decision and does not affect the variables of interest (adult and child labour) directly.

In this paper, I use the perceptions of erosion as an instrument. In theory, this variable should be highly correlated with the adoption of Soil and Water conservation technologies. Specifically, a negative correlation[17] is expected since, soil and water conservation technologies are more likely to be adopted in plots where erosion is perceived to be a problem. However, conceptually, it is also possible that erosion perceptions influence the amount of labour directly.

[17] The variable is labelled as follows. 1 means that the household does not perceive erosion as being a problem for that particular plot. 0 means that the household perceives erosion to be a problem. As a result, we would expect households to adopt SWC technologies in plots where they believe erosion is a problem (when the variable takes a value of 0).

It is possible that the level of effort exerted by households depends on the level of perceived erosion in a plot, although the direction is ambiguous. Specifically, households may work more in a plot where erosion is a problem in order to compensate for the negative effects of erosion, in which case it would lead to an increase in labour. Alternatively, they may decide to exert more effort in other less-eroded plots or simply increasingly shift away from agriculture, in which case a negative direct effect of labour would be expected. The existence of a direct effect may be particularly problematic in this paper since only one instrument is used, which implies that I cannot test for the exogeneity of the instrument.

Therefore, in order to substantiate the admissibility of the instrument used, I proceed to the same falsification test proposed in Di Falco et al. (2011). The test consists of running a probit regression to show that the instrument is highly correlated with the adoption decision, then run an OLS regression on the dependent variables for non-adopters. If the instrument has a statistically significant effect on adoption and a statistically insignificant effect on the levels of child and adult labour of non-adopters, this suggests that this instrument may be acceptable.

The results of the falsification test are presented in Table 4.2. The results show a strong negative relationship between the instrument and the adoption equation. The associated t-statistics (in italic) for the instrument are approximately 12 for adult labour and above 8[18] for the sub-sample of plots that use child labour. As such, this instrument qualifies as a strong predictor of the adoption decision. Regarding the coefficient of the instrument in the labour equations for the non-adopters, its statistical significance is comfortably rejected at the usual significance levels. As such, this instrument satisfies the conditions laid by Di Falco et al. (2011). In Appendix C I discuss further why I did not use alternative instruments. Typically, other instruments proposed in the literature either 1) failed the falsification test; 2) were not strong predictors of the adoption decision; 3) displayed different signs across different sub-samples; and/or 4) were incompatible with either Kebele or Household fixed effects as data was collected either at the Kebele or household level.

[18]It could be argued that a t-statistic of 8 is not powerful enough for a strong instrument. As such, a weak instrument p-value was computed in all of our child labour IV results.

Table 4.2: Falsification tests

	Adult labour			Child labour		
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Variables	Adoption	OLS	OLS	Adoption	OLS	OLS
Distance to market (km)	0.004	0.028***	-0.034	-0.001	0.007	-0.003
	(0.015)	(0.011)	(0.082)	(0.018)	(0.017)	(0.027)
Hailstorm since 1994	-0.012	0.167	-1.032**	-0.31	0.092	-0.932**
	(0.159)	(0.121)	(0.399)	(0.209)	(0.285)	(0.358)
Has TV (1 if yes, 0 if no)	-0.204	0.13	0.88	-0.471*	0.138	-0.198
	(0.171)	(0.139)	(0.570)	(0.242)	(0.216)	(0.458)
Is head of household literate (1 Yes)	-0.067	0.128*	1.532***	0.242	-0.174	0.832
	(0.117)	(0.072)	(0.338)	(0.168)	(0.173)	(1.109)
At least one flood since 1994	0.013	-0.022	-1.292***	-0.262	-0.193	-0.207
	(0.171)	(0.115)	(0.373)	(0.201)	(0.241)	(0.430)
At least on drought since 1994	0.027	-0.003	-1.449	-0.182	0.143	0.545
	(0.145)	(0.164)	(1.530)	(0.180)	(0.237)	(0.351)
Plot is highly fertile	0.233**	0.004	-0.138	0.288*	-0.144	-0.411*
	(0.101)	(0.072)	(0.122)	(0.155)	(0.129)	(0.238)
Plot is steep	0.520**	-0.111	0.864	0.716**	0.051	-0.041
	(0.227)	(0.251)	(0.537)	(0.340)	(0.351)	(0.313)
Ln Area plot (ha)	0.07	0.209***	0.141*	-0.04	0.041	0.106
	(0.058)	(0.047)	(0.084)	(0.078)	(0.090)	(0.120)
Plot is flat	0.132	-0.036	0.297	-0.242	0.113	-0.041
	(0.109)	(0.077)	(0.180)	(0.162)	(0.152)	(0.313)
Plot has medium depth	0.336***	0.04	-0.043	0.310**	0.083	-0.372
	(0.096)	(0.070)	(0.272)	(0.155)	(0.145)	(0.427)
Household has child under 6	-0.236*	-0.035	-0.161	-0.306*	-0.619***	-0.57
	(0.132)	(0.089)	(0.841)	(0.170)	(0.178)	(0.345)
ln Number of children under 6	0.089	-0.02	-0.06	0.151	-0.746***	-0.522***
	(0.168)	(0.113)	(0.431)	(0.237)	(0.267)	(0.183)
Household has child aged 6-11	-0.124	0.002	-0.889**	-0.108	0.077	-0.63
	(0.134)	(0.085)	(0.399)	(0.200)	(0.192)	(0.545)
ln Number of children aged 6-11	-0.065	-0.045	-0.525	0.028	0.092	0.526
	(0.146)	(0.084)	(0.458)	(0.217)	(0.170)	(0.350)
Household has child aged 11-15	0.04	-0.067	-2.262***	0.237	-0.058	-0.618
	(0.131)	(0.086)	(0.216)	(0.232)	(0.205)	(0.404)
ln Number of children aged 11-15	-0.178	-0.047	-1.924**	0.041	0.343	0.059
	(0.159)	(0.113)	(0.908)	(0.217)	(0.207)	(0.588)
ln Number of plots	0.08	-0.056	-2.181	0.023	0.09	-1.453*
	(0.135)	(0.091)	(1.429)	(0.224)	(0.205)	(0.830)
ln Number of adults	0.031	0.203**	2.103***	0.437**	0.007	2.017***
	(0.134)	(0.100)	(0.580)	(0.195)	(0.270)	(0.374)
No perceived erosion on plot (1-yes, 0 no)	-1.370***	-0.096	-0.265	-1.447***	0.153	-0.414
	(0.113)	(0.105)	(0.212)	(0.176)	(0.186)	(0.427)
	-12.11	*-0.913*	*-1.249*	*-8.241*	*0.825*	*-0.969*
Constant	-0.518	2.874***	6.125***	-1.330*	0.271	3.797**
	(0.468)	(0.297)	(1.449)	(0.780)	(0.574)	(1.734)
Kebele FE	✓	✓		✓	✓	
Household FE			✓			✓
Number of observations	3447	861	269	1512	306	132

Number in parentheses denote standard errors. In the case of our instrument, the numbers in *italic* denote t-values. *, **, *** denote statistical significance at the 10%, 5% and 1% level, respectively.
Note: Columns (2), (3), (5) and (6) only use the sub-sample of non-adopters, whereas Columns (1) and (4) use the full sample.

4.6 Results

Table 4.3 shows the output of the endogenous switching regression models for both child labour and adult labour[19][20].

The child labour results in table 4.3 suggest that plot depth and fertility are significant drivers of Soil and Water Conservation technology adoption. Demographic characteristics also seem to play a role in the adoption of SWC technologies. Specifically, households with a larger number of adults are more likely to adopt. A similar positive (though statistically insignificant) coefficient is associated with the number of children aged 11-15. However, a higher number of children under 6 is associated with a lower probability of adoption. This suggests that adoption is correlated with the amount of labour availability, as older children and teenagers may be more able to work compared to young children.

Previous exposure to natural disasters is also negatively correlated with adoption. However, this is only significant in the case of a hail storm. Households who own a TV are less likely to adopt. Finally, we also find that households in more remote areas are less likely to adopt SWC technologies than those closer to the market, although this is not statistically significant. As expected, our identifying variable (erosion perceptions) is strongly (t-stat=10.72) negatively correlated with adoption.

However, the estimated rhos are both statistically insignificant. As a result there is no strong evidence supporting the existence of self-selection in the child labour regressions. Nevertheless, the sign of the rhos indicates that plots where SWC technologies are used tend to use more child labour than a random plot.

The results displayed in Table 4.3 also highlight the difference in terms of some of the coefficients the determinants of the level of child labour use. Households with large numbers of

[19]The results for child labour refer to the sub-sample of plots where child labour is used for villages which have enough variation of adoption.

[20]In both cases, I only kept Kebeles with at least 5 observations in each adoption status. This means that a number of observations are not used, especially in the child labour specification. In total, in the child labour specification, 4 Kebeles are not used (196 observations). However, in the robustness checks using an Instrumental variable approach, I test the sensitivity to the inclusion/exclusion of these Kebeles. The main reason driving this choice in the ESRM specification is that I believe that the inclusion of village fixed effects is key to capture important features driving the outcome, including weather, labour market arrangements and agro-ecological suitability.

Table 4.3: Endogenous Switching Regression Model Results

	Child Labour			Adult Labour		
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Variables	Non-Adopters	Adopters	Selection	Non-Adopters	Adopters	Selection
Distance to market (km)	-0.003	0.003	-0.008	0.028***	-0.001	0.002
	(0.014)	(0.006)	(0.013)	(0.008)	(0.003)	(0.010)
Hailstorm since 1994	0.042	0.016	-0.322**	0.171**	0.083**	-0.024
	(0.185)	(0.072)	(0.142)	(0.087)	(0.032)	(0.095)
Has TV (1 if yes, 0 if no)	0.091	0.006	-0.461***	0.139*	0.146***	-0.200**
	(0.170)	(0.083)	(0.162)	(0.078)	(0.035)	(0.097)
Is head of household literate (1 Yes)	-0.176	-0.171***	0.240**	0.130***	-0.032	-0.046
	(0.115)	(0.063)	(0.116)	(0.047)	(0.025)	(0.067)
At least one flood since 1994	-0.18	0.021	-0.236	-0.024	-0.063*	0.019
	(0.153)	(0.072)	(0.144)	(0.094)	(0.033)	(0.100)
At least on drought since 1994	0.101	-0.246***	-0.235*	-0.004	0.101***	0.013
	(0.183)	(0.066)	(0.136)	(0.090)	(0.028)	(0.088)
Plot is highly fertile	-0.072	-0.104	0.267**	-0.012	0.032	0.266***
	(0.109)	(0.068)	(0.116)	(0.050)	(0.027)	(0.071)
Plot is steep	0.087	-0.18*	0.707*	-0.141	0.013	0.555***
	(0.461)	(0.105)	(0.375)	(0.155)	(0.049)	(0.196)
Ln Area plot (ha)	0.071	0.138***	-0.011	0.209***	0.228***	0.065
	(0.074)	(0.039)	(0.080)	(0.035)	(0.018)	(0.050)
Plot is flat	0.116	0.029	-0.213	-0.046	0.016	0.116
	(0.140)	(0.072)	(0.144)	(0.058)	(0.027)	(0.083)
Plot has medium depth	0.099	0.113*	0.329***	0.021	0.108***	0.342***
	(0.121)	(0.061)	(0.121)	(0.053)	(0.025)	(0.069)
Household has child under 6	-0.532***	-0.083	-0.316**	-0.025	-0.003	-0.237***
	(0.138)	(0.070)	(0.127)	(0.061)	(0.029)	(0.079)
ln Number of children under 6	-0.639***	-0.059	0.124	-0.023	-0.021	0.086
	(0.179)	(0.086)	(0.174)	(0.070)	(0.035)	(0.098)
Household has child aged 6-11	0.076	-0.127*	-0.129	0.005	-0.047	-0.1
	(0.136)	(0.076)	(0.140)	(0.059)	(0.030)	(0.081)
ln Number of children aged 6-11	0.082	0.12	0.024	-0.046	-0.038	-0.057
	(0.124)	(0.078)	(0.140)	(0.057)	(0.033)	(0.088)
Household has child aged 11-15	0.013	-0.210**	0.209	-0.071	-0.055*	0.043
	(0.142)	(0.082)	(0.152)	(0.056)	(0.029)	(0.078)
ln Number of children aged 11-15	0.338**	0.327***	0.038	-0.036	-0.007	-0.190**
	(0.142)	(0.075)	(0.141)	(0.068)	(0.036)	(0.096)
ln Number of plots	0.087	-0.145*	0.037	-0.06	-0.111***	0.051
	(0.155)	(0.082)	(0.154)	(0.062)	(0.034)	(0.088)
ln Number of adults	0.139	0.155**	0.408***	0.198***	0.278***	0.064
	(0.177)	(0.076)	(0.146)	(0.064)	(0.031)	(0.082)
No perceived erosion on plot (1-yes, 0 no)			-1.432***			-1.362***
			(0.134)			(0.080)
Constant	2.139***	2.817***	0.491	2.932***	3.471***	1.093***
	(0.386)	(0.281)	(0.420)	(0.184)	(0.097)	(0.233)
Kebele Fixed Effects	✓	✓	✓	✓	✓	✓
			Ancillary Parameters			
ln	-0.339***	-0.102***		-0.497***	-0.586***	
	(0.042)	(0.025)		(0.026)	(0.015)	
Rho	0.037	-0.35		-0.164	-0.153	
	(0.170)	(0.221)		(0.115)	(0.100)	
p-value LR test		0.329			0.108	
Number of observations	289	1169	1458	857	2456	3313

Number in parentheses denote standard errors. In the case of our instrument, the numbers in *italic* denote t-values. *, **, *** denote statistical significance at the 10%, 5% and 1% level, respectively.

children under the age of 6 tend to use less child labour, whereas those with a larger numbers of children between the ages of 11 and 15 tend to use larger amounts of child labour. This suggests that older children are likely to be preferred to carry out farm labour.

Impact estimates in Table 4.4 shows the ATE estimates for child labour. The results suggest that adopting a SWC technology leads to an increase of approximately 29% per plot. In levels, this corresponds to approximately 1.26 extra days per plot. However, the ATT estimate is much lower than the ATU. The ATT estimate suggests a 14% increase in child labour for adopters (0.46 additional days per plot), whereas the ATU estimate suggests an increase in child labour of 117% (4.5 additional days per plot)[21].

Table 4.4: Treatment Effects

	ATE				ATT				ATU			
Variable	N	Mean	S.E.	t-value	N	Mean	S.E.	t-value	N	Mean	S.E.	t-value
Child Labour (logs)	1458	0.258***	0.019	13.459	1169	0.13***	0.020883	6.236	289	0.774***	0.032647	23.724
Child Levels (levels)	1458	1.256***	0.10748	11.684	1169	0.457***	0.105263	4.346	289	4.484***	0.260529	17.212
Adult Labour (logs)	3313	0.273***	0.004726	57.885	2456	0.3***	0.005509	54.455	857	0.197***	0.008676	22.755
Adult Labour (levels)	3313	4.845***	0.108	44.872	2456	5.175***	0.1256	41.201	857	3.900***	0.208	18.737

*, **, *** denote statistical significance at the 10%, 5% and 1% level, respectively. N refers to the total number of observations. S.E. refers to the standard error.

These results suggest that adoption of SWC technologies leads to an increase in the amount of child labour used. In addition to this, estimated impacts are higher for non-adopters than for adopters. One possible explanation for this is that, on average, adults already work more days in households that do not use SWC technologies in some of their plots. As a result, children may be expected to take on more of the additional labour generated by the adoption of the new technology.

We then test the robustness of the result using a set of standard IV regressions using four different specifications. First, I use the full sample including the villages excluded for the purposes of the ESRM. In the second specification, I use the same sample used in the ESRM specification. In the third specification, I use a sample of partial adopters (those who adopted SWC technologies in some, but not all, plots). Finally, in the final specification, I use the sub-sample of partial adopters and control for household fixed-effects. In the last specification, however, owing to the inclusion of fixed effects, I only use plot-level variables[22] . The results

[21]Note that the (log) results in Table 4.4 refers to the log difference. As a result, to obtain the percentage change, we need to use the following formula $(e^{iestimate} - 1) * 100$, where iestimate refers to the impact estimate.

[22]Given that the dataset is a cross-section, household level variables are, by definition, invariant at the household level and thus cannot be included alongside fixed effects.

are summarized in table 4.5 and the estimated magnitude of the estimated impacts is very similar to the impacts estimated in the ESRM regression and range from 30% to 34%. In terms of significance, however, the impacts are no longer significant at the conventional levels of statistical significance.

Table 4.5: Robustness Checks: IV regressions Child Labour

	All		Exclude Keb 1, 8 and 10		FE sample		FE	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Variables	C. labour	Adoption	C. labour	Adoption	C. labour	Adoption	C. labour	Adoption
Adopted SWC practices	0.294		0.263		0.275		0.287	
	(0.264)		(0.256)		(0.236)		(0.204)	
Distance to market (km)	0.005	0.000	0.002	-0.001	0.005	-0.005		
	(0.007)	(0.002)	(0.007)	(0.002)	(0.029)	(0.005)		
Hailstorm since 1994	0.004	-0.053	0.014	-0.056	-0.140	-0.168***		
	(0.117)	(0.033)	(0.127)	(0.036)	(0.394)	(0.051)		
Has TV (1 if yes, 0 if no)	-0.018	-0.057	0.014	-0.072*	-0.158	-0.081		
	(0.114)	(0.037)	(0.116)	(0.039)	(0.292)	(0.091)		
Is head of household literate (1 Yes)	-0.102	0.034	-0.180**	0.035	0.003	-0.063		
	(0.091)	(0.027)	(0.088)	(0.031)	(0.221)	(0.050)		
At least one flood since 1994	0.017	-0.040	0.017	-0.045	0.167	-0.029		
	(0.109)	(0.033)	(0.112)	(0.034)	(0.241)	(0.044)		
At least on drought since 1994	-0.157	-0.026	-0.236**	-0.037	-0.050	-0.225***		
	(0.106)	(0.026)	(0.110)	(0.029)	(0.269)	(0.056)		
Plot is highly fertile	-0.111	0.049*	-0.116	0.050	-0.172	-0.040	-0.007	-0.066
	(0.081)	(0.029)	(0.076)	(0.031)	(0.129)	(0.075)	(0.072)	(0.111)
Plot is steep	-0.131	0.062**	-0.120	0.048	0.170	-0.001	-0.101	-0.132
	(0.128)	(0.030)	(0.133)	(0.031)	(0.225)	(0.083)	(0.205)	(0.089)
Ln Area plot (ha)	0.115***	-0.004	0.130***	-0.004	0.134**	0.007	0.181**	-0.022
	(0.041)	(0.012)	(0.043)	(0.012)	(0.066)	(0.032)	(0.074)	(0.043)
Plot is flat	-0.028	-0.034*	-0.010	-0.041*	0.107	-0.110	0.003	-0.089
	(0.073)	(0.020)	(0.076)	(0.023)	(0.170)	(0.071)	(0.125)	(0.088)
Plot has medium depth	0.134*	0.031	0.119	0.040	0.105	0.030	-0.126	-0.057
	(0.073)	(0.022)	(0.079)	(0.025)	(0.156)	(0.058)	(0.134)	(0.094)
Household has child under 6	-0.172*	-0.048	-0.201**	-0.063*	-0.648***	-0.124***		
	(0.096)	(0.030)	(0.102)	(0.033)	(0.217)	(0.044)		
ln Number of children under 6	-0.191	0.012	-0.166	0.000	0.299	0.015		
	(0.128)	(0.038)	(0.131)	(0.040)	(0.339)	(0.068)		
Household has child aged 6-11	-0.128	-0.019	-0.111	-0.018	0.036	0.000		
	(0.102)	(0.033)	(0.105)	(0.036)	(0.278)	(0.074)		
ln Number of children aged 6-11	0.082	0.003	0.113	0.002	-0.296	0.020		
	(0.116)	(0.036)	(0.113)	(0.039)	(0.313)	(0.072)		
Household has child aged 11-15	-0.151	0.045	-0.115	0.054	-0.518	0.014		
	(0.110)	(0.039)	(0.116)	(0.042)	(0.317)	(0.075)		
ln Number of children aged 11-15	0.316***	0.012	0.372***	0.013	0.103	-0.136**		
	(0.113)	(0.035)	(0.116)	(0.038)	(0.299)	(0.052)		
ln Number of plots	-0.198	-0.008	-0.139	-0.008	-0.511	0.148**		
	(0.121)	(0.034)	(0.117)	(0.038)	(0.316)	(0.073)		
ln Number of adults	0.287**	0.074**	0.195*	0.091***	0.454	-0.056		
	(0.116)	(0.031)	(0.115)	(0.035)	(0.336)	(0.065)		
No perceived erosion on plot (1-yes, 0 no)		-0.261***		-0.284***		-0.552***		-0.569***
		(0.031)		(0.033)		(0.075)		(0.082)
Constant	0.996**	0.219*	1.000**	0.752***	3.232***	1.030***	2.218***	1.124***
	(0.318)	(0.087)	(0.308)	(0.094)	(0.650)	(0.205)	(0.221)	(0.125)
Kebele FE	✓		✓		✓			
Household FE							✓	
Number of plots	1654	1654	1472	1472	312	312	312	362
Number of Households	502		439		78		78	
R-squared	0.237	0.415	0.229	0.415	0.407	0.486	0.758	0.48
Weak iv p-value	0.264		0.304		0.241		0.186	

Number in parentheses denote standard errors. *, **, *** denote statistical significance at the 10%, 5% and 1% level, respectively. C. labour refers to the total days of child labour in the household.

In addition to the results in Tables 4.3 and 4.5, I also test whether the adoption of Soil and Water Conservation technologies has an impact on the probability of using child labour. The results, summarized in Table 4.6, are similar to the other results using IV and suggest a positive but insignificant coefficient. As such, I find no strong evidence that adoption of Soil and Water Conservation technologies positively affects the probability of using of child labour.

The output of the ESRM regressions for family labour are summarized in table 4.3. With regards to the adoption equation (column 6), the results are qualitatively similar to that of the child labour selection equation. Certain plot characteristics, such as depth, inclination and fertility remain highly significant, whereas the presence of children under 6 and the ownership of a TV are negatively correlated with the adoption decision. The identifying variable is, once again, negatively related to the adoption decision and highly significant (t-statistic of approximately 17). In the case of family labour, the estimated rho is almost significant at the 10% level.

The regression results indicates that, for non-adopters, distance to market is positively and significantly related to the use of family labour. Similarly, literate household heads tend to use more adult family labour. Finally, as would be expected, the area of the plot as well as the total number of adults in the household positively affects the amount of labour used.

Table 4.4 shows the estimated impacts for family labour. Overall, there is strong evidence suggesting that the adoption of SWC technologies leads to a large increase in adult family labour at the plot-level. The overall predicted increase in household adult labour is of around 31% per plot, with the ATT (35%) being larger than the ATU(21%). This result is, to some extent, surprising as we would expect those with the lower cost of adoption (i.e. lower increases in labour) to adopt. However, in our context, the result makes sense. The differences between the ATT and the ATU could be explained by the fact that adults in households that have plots where they have adopted SWC technologies work fewer hours.

As a result, adults are able to increase the number of days worked as a result of adopting SWC technologies. Non-adopters, on the other hand, already work more days. As a result, they may be less inclined to adopt labour-intensive technologies, even if they have lower labour costs of adoption. In addition to this, these results also highlight that the assumption of no changes in inputs as a result of the adoption process, at least in this case, does not seem to be plausible.

Table 4.6: IV results: Probability to use Child Labour

	(1)	(2)
Variables	Child Labour	Adoption
Adopted SWC practices	0.111	
	(0.312)	
Distance to market (km)	0.021*	0.001
	(0.011)	(0.002)
Hailstorm since 1994	-0.175	-0.002
	(0.138)	(0.029)
Has TV (1 if yes, 0 if no)	-0.219	-0.017
	(0.138)	(0.031)
Is head of household literate (1 Yes)	-0.046	-0.009
	(0.095)	(0.021)
At least one flood since 1994	0.334**	0.002
	(0.145)	(0.029)
At least on drought since 1994	-0.165	0.011
	(0.111)	(0.023)
Plot is highly fertile	0.049	0.051**
	(0.089)	(0.020)
Plot is steep	0.182	0.073**
	(0.151)	(0.029)
Ln Area plot (ha)	0.02	0.011
	(0.045)	(0.01)
Plot is flat	-0.003	0.008
	(0.08)	(0.017)
Plot has medium depth	0.065	0.055***
	(0.079)	(0.017)
Household has child under 6	-0.052	-0.035
	(0.112)	(0.025)
ln Number of children under 6	-0.296**	0.015
	(0.135)	(0.030)
Household has child aged 6-11	-0.327***	-0.019
	(0.116)	(0.026)
ln Number of children aged 6-11	0.085	-0.008
	(0.123)	(0.027)
Household has child aged 11-15	-0.789***	0.005
	(0.107)	(0.025)
ln Number of children aged 11-15	0.306**	-0.034
	(0.139)	(0.030)
ln Number of plots	0.044	0.01
	(0.117)	(0.026)
ln Number of adults	0.235**	0.01
	(0.114)	(0.026)
No perceived erosion on plot (1-yes, 0 no)		-0.261***
		(0.022)
Constant		0.273***
		(0.086)
Kebele FE	✓	✓
Number of observations	3651	

Number in parentheses denote standard errors. *, **, *** denote statistical significance at the 10%, 5% and 1% level, respectively.

I perform the same robustness checks using an set of IV regressions and these are presented in Table 4.7. Unlike in the case of child labour, the results remain large (and similar to the ESRM estimates) and statistically significant at least at the 5% level. Across specifications, the estimated average impacts range from approximately 25% to 36%).

Table 4.7: Robustness Checks: IV regressions Adult Labour

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	All		Exclude Keb 1 and 10		FE sample		FE	
Variables	A. Labour	Adoption	A. Labour	Adoption	A. Labour	Adoption	A. Labour	Adoption
Adopted SWC practices	0.254**		0.269**		0.309***		0.223***	
	(0.116)		(0.114)		(0.095)		(0.075)	
Distance to market (km)	0.003	0.001	0.004	0.001	0.010	-0.005		
	(0.004)	(0.002)	(0.004)	(0.002)	(0.007)	(0.004)		
Hailstorm since 1994	0.070	-0.002	0.095*	-0.002	0.184	-0.009		
	(0.047)	(0.029)	(0.049)	(0.031)	(0.153)	(0.061)		
Has TV (1 if yes, 0 if no)	0.158***	-0.016	0.160***	-0.028	0.348***	-0.146**		
	(0.050)	(0.031)	(0.052)	(0.033)	(0.118)	(0.058)		
Is head of household literate (1 Yes)	0.010	-0.009	0.005	-0.011	0.087	-0.057		
	(0.031)	(0.021)	(0.033)	(0.023)	(0.068)	(0.046)		
At least one flood since 1994	-0.064	0.002	-0.067	0.002	-0.214	0.012		
	(0.049)	(0.029)	(0.050)	(0.030)	(0.141)	(0.058)		
At least on drought since 1994	0.106***	0.010	0.102**	0.008	0.142	-0.078		
	(0.039)	(0.023)	(0.043)	(0.026)	(0.093)	(0.056)		
Plot is highly fertile	0.020	0.053**	0.019	0.062***	0.006	-0.053	0.038	-0.091
	(0.031)	(0.021)	(0.032)	(0.021)	(0.061)	(0.049)	(0.047)	(0.065)
Plot is steep	-0.034	0.071**	-0.025	0.063**	0.097	-0.081	0.094	-0.147
	(0.064)	(0.029)	(0.061)	(0.030)	(0.126)	(0.081)	(0.133)	(0.105)
Ln Area plot (ha)	0.219***	0.010	0.225***	0.007	0.275***	0.009	0.260***	0.009
	(0.022)	(0.010)	(0.024)	(0.011)	(0.044)	(0.024)	(0.046)	(0.031)
Plot is flat	0.007	0.009	-0.003	0.006	0.021	-0.011	0.104**	-0.02
	(0.030)	(0.018)	(0.031)	(0.018)	(0.060)	(0.047)	(0.050)	(0.070)
Plot has medium depth	0.067**	0.056***	0.086***	0.062***	0.121*	0.073*	-0.089	0.008
	(0.031)	(0.017)	(0.033)	(0.019)	(0.064)	(0.042)	(0.059)	(0.065)
Household has child under 6	0.005	-0.036	0.002	-0.042	-0.054	-0.081		
	(0.038)	(0.025)	(0.040)	(0.027)	(0.100)	(0.057)		
ln Number of children under 6	-0.010	0.015	-0.021	0.010	-0.274**	0.011		
	(0.047)	(0.030)	(0.050)	(0.033)	(0.116)	(0.060)		
Household has child aged 6-11	-0.030	-0.019	-0.026	-0.019	-0.056	0.043		
	(0.039)	(0.026)	(0.042)	(0.028)	(0.083)	(0.059)		
ln Number of children aged 6-11	-0.037	-0.007	-0.042	-0.008	-0.044	-0.001		
	(0.042)	(0.027)	(0.045)	(0.029)	(0.124)	(0.062)		
Household has child aged 11-15	-0.051	0.005	-0.056	0.009	-0.064	-0.004		
	(0.038)	(0.025)	(0.040)	(0.027)	(0.098)	(0.064)		
ln Number of children aged 11-15	-0.014	-0.033	-0.005	-0.034	-0.098	-0.021		
	(0.050)	(0.030)	(0.053)	(0.032)	(0.134)	(0.060)		
ln Number of plots	-0.070*	0.009	-0.086*	0.004	-0.083	0.126*		
	(0.041)	(0.026)	(0.044)	(0.029)	(0.109)	(0.068)		
ln Number of adults	0.253***	0.010	0.259***	0.017	0.267***	-0.030		
	(0.049)	(0.026)	(0.053)	(0.029)	(0.086)	(0.056)		
No perceived erosion on plot (1-yes, 0 no)		-0.261***		-0.281***		-0.543***		-0.657***
		(0.022)		(0.024)		(0.055)		(0.063)
Constant	2.961***	0.772***	2.985***	0.785***	2.594***	0.917***	2.402***	0.972***
	(0.139)	(0.079)	(0.145)	(0.082)	(0.394)	(0.195)	(0.116)	(0.101)
Kebele FE	✓		✓		✓			
Household FE							✓	
Number of observations	3630	3630	3313	3313	685	685	685	685
Number of Households	969		874		151		151	
R-squared	0.286	0.441	0.289	0.434	0.337	0.3625	0.672	0.456

Number in parentheses denote standard errors. *, **, *** denote statistical significance at the 10%, 5% and 1% level, respectively. A. labour refers to the total number of days worked on-farm by adults in the household.

4.7 Conclusion

Overall, the results clearly suggest an increase in adult family labour. In the case of child labour, all estimates suggest a positive and large coefficient, although this is only significant in the case of the endogenous switching regression model.

These results are meaningful for both policy-makers and research for a number of reasons. First, the large impacts on labour suggest that, at least in our case, using methods such as PSM or ESRM, which assume no changes in inputs as a consequence of adoption, may be problematic. Therefore, not taking into account the impacts of technology adoption on input-use may have large effects in the impact estimates on output. However, I am unable to speculate on how often these occur, how important this omission might be. Moreover, I do not propose an alternative way to deal with this issue.

Secondly, the results on labour also suggest that, at least in this case, widespread adoption of SWC technologies could potentially have unintended spill-over effects in terms of child labour. If these effects are well understood, policy-makers may also be able to devise schemes and policies in such a way as to mitigate any undesirable negative spill-overs. Finally, while remaining speculative, the results presented suggest a potential channel to explain the puzzling result in Di Falco et al. (2011), who found that non-adopters had the highest predicted gains from adoption. I find that adults in households that have plots where SWC technologies have not been adopted already work more days. As a result, non-adopters display higher impacts on child labour, as children may be expected to take on more of the additional labour caused by the adoption of a SWC technology. Consequently, non-adopters may be more reluctant to adopt because 1) they already work longer hours; or 2) because it would require a large increase in child labour. These two channels would provide a plausible explanation to why, despite higher gains, non-adopters decided not to adopt.

References

[1] Bachewe, F. N., Berhane, G., Minten, B., & Taffesse, A. S. (2016). Non-farm income and labor markets in rural Ethiopia (Vol. 90). *International Food Policy Research Institute*.

[2] Basu, A., Heckman, J. J., Navarro-Lozano, S. & Urzua, S. (2007). Use of instrumental variables in the presence of heterogeneity and self-selection: An application to breast Cancer patients. *Health Economics*, 16(11), 1133-57.

[3] Beegle, K., Dehejia, R., & Gatti, R. (2009). Why should we care about child labor? The education, labor market, and health consequences of child labor. *Journal of Human Resources*, 44(4), 871-889.

[4] Bekele, W., & Drake, L. (2003). Soil and water conservation decision behavior of subsistence farmers in the Eastern Highlands of Ethiopia: a case study of the Hunde-Lafto area. *Ecological economics*, 46(3), 437-451.

[5] Bewket, W. (2007). Soil and water conservation intervention with conventional technologies in northwestern highlands of Ethiopia: Acceptance and adoption by farmers. *Land use policy*, 24(2), 404-416.

[6] Byerlee, D., De Janvry, A., & Sadoulet, E. (2009). Agriculture for development: Toward a new paradigm. *Annual Review of Resource Economics*, 1, 15-31.

[7] Collier, P., & Dercon, S. (2014). African Agriculture in 50 Years: Smallholders in a Rapidly Changing World? *World Development*, 63, 92-101.

[8] Colmer, J. (2013). Climate Variability, Child Labour and Schooling: Evidence on the Intensive and Extensive Margin.

[9] Deressa, T. T., Hassan, R. M., Ringler, C., Alemu, T., & Yesuf, M. (2009). Determinants

of farmers choice of adaptation methods to climate change in the Nile Basin of Ethiopia. *Global environmental change*, 19(2), 248-255.

[10] Diao, X., Hazell, P., & Thurlow, J. (2010). The role of agriculture in African development. *World development*, 38(10), 1375-1383.

[11] Dillon, B., & Barrett, C. B. (2017). Agricultural factor markets in Sub-Saharan Africa: an updated view with formal tests for market failure. *Food policy*, 67, 64-77.

[12] Di Falco, S., Veronesi, M., & Yesuf, M. (2011). Does adaptation to climate change provide food security? A micro-perspective from Ethiopia. *American Journal of Agricultural Economics*, aar006

[13] Duflo, E., Glennerster, R. & Kremer, M. (2007). Using Randomization in Development Economics Research: A Toolkit. *Handbook of Development Economics*, 4, 3895-3962.

[14] Fernandez-Cornejo, J., Hendricks, C. & Mishra, A. K. (2005). Technology Adoption and Off-Farm Household Income: The Case of Herbicide-Tolerant Soybeans. *Journal of Agricultural and Applied Economics*, 37(3), 549-563.

[15] Fisher, R. A. (1935) The design of Experiments. Edinburgh: Oliver and Boyd, 1935.

[16] Godfray, H. C. J., Beddington, J. R., Crute, I. R., Haddad, L., Lawrence, D., Muir, J. F., Pretty, J., Robinson, S., Thomas, S. M. & Toulmin, C. (2010). Food security: the challenge of feeding 9 billion people. *science*, 327(5967), 812-818.

[17] Gorst, A., Groom, B., & Dehlavi, A. (2015). Crop productivity and adaptation to climate change in Pakistan. *Grantham Research Institute on Climate Change and the Environment Working paper*, (189).

[18] Giller, K. E., Witter, E., Corbeels, M., & Tittonell, P. (2009). Conservation agriculture and smallholder farming in Africa: the heretics view. *Field crops research*, 114(1), 23-34.

[19] Hawkes, C., & Ruel, M. (2006). The links between agriculture and health: an intersectoral opportunity to improve the health and livelihoods of the poor. *Bulletin of the World Health Organization*, 84(12), 984-990.

[20] Heady, C. (2000). What is the effect of child labour on learning achievement? Evidence from Ghana (*Innocenti Working Papers*, no. 79).

[21] Heckman, J. J. & Vytlacil, E. J. (2007) Chapter 71 Econometric Evaluation of Social Programs, Part II: The Marginal Treatment Effect to Organize Alternative Econometric Estimators to Evaluate Social Programs, and to Forecast their Effects in New Environments. *Handbook of Econometrics*, 6(B), 4875-5143

[22] Heckman, J. J. & Robb, , R. (1985). Alternative methods for evaluating the impact of interventions: An overview. *Journal of Econometrics*, 30(1-2), 239-267

[23] Huffman, W. E. (1991). Agricultural household models: survey and critique. *Multiple job-holding among farm families*, 79-111.

[24] Imbens, G. W. & Angrist, J. D. (1994). Identification and estimation of local average treatment effects". *Econometrica*, 62(2), 467-475.

[25] Kassie, M., Pender, J., Yesuf, M., Kohlin, G., Bluffstone, R., & Mulugeta, E. (2008). Estimating returns to soil conservation adoption in the northern Ethiopian highlands. Agricultural Economics, 38(2), 213-232.

[26] Kato, E., Ringler, C., Yesuf, M., & Bryan, E. (2011). Soil and water conservation technologies: a buffer against production risk in the face of climate change? Insights from the Nile basin in Ethiopia. *Agricultural Economics*, 42(5), 593-604.

[27] Kassie, M., Khlin, G., Bluffstone, R., & Holden, S. (2011). Are soil conservation technologies winwin? A case study of Anjeni in the northwestern Ethiopian highlands. In *Natural Resources Forum* (Vol. 35, No. 2, pp. 89-99). Blackwell Publishing Ltd.

[28] Kassie, M., Teklewolde, H., & Erenstein, O. (2015). Adoption Pathways project discussion paper 2 February 2015. *Technology.*

[29] Khonje, M., Manda, J., Alene, A. D., & Kassie, M. (2015). Analysis of adoption and impacts of improved maize varieties in eastern Zambia. *World Development*, 66, 695-706.

[30] Larney, F. J., Olson, B. M., Janzen, H. H., & Lindwall, C. W. (2000). Early impact of topsoil removal and soil amendments on crop productivity. *Agronomy journal*, 92(5), 948-956.

[31] Lele, U., & Mellor, J. W. (1981). Technological change, distributive bias and labor transfer in a two sector economy. *Oxford Economic Papers*, 33(3), 426-441.

[32] Lewis, W. A. (1954). Economic development with unlimited supplies of labour. *The manchester school*, 22(2), 139-191.

[33] McCarthy, N. (2011). Understanding agricultural households' adaptation to climate change and implications for mitigation: land management and investment options.

[34] Maddala, G. S. (1983). Models with self-selectivity. *Limited-Dependent and Qualitative Variables in Economics*, 3, 257-291.

[35] Mendelsohn, R. (2008). The impact of climate change on agriculture in developing countries. *Journal of Natural Resources Policy Research*, 1(1), 5-19.

[36] Agricultural Technology Adoption and Poverty Reduction: A Propensity-Score Matching Analysis for rural Bangladesh. *Food Policy*, 32, 372-393

[37] Haile, M., Herweg, K. G., and Stillhardt, B. (2006). Sustainable land management: A new approach to soil and water conservation in Ethiopia.

[38] Neyman, J. (1990). On the Application of Probability Theory to Agricultural Experiments. Essay on Principles. Section 9. *Statistical Science*, 5(4) 465-472. Translated by Dabrowska, D. M. and Speed, P. T.

[39] Osman, M., & Sauerborn, P. (2001). Soil and water conservation in Ethiopia. *Journal of soils and sediments*, 1(2), 117-123.

[40] Palacios-Lopez, A., Christiaensen, L., & Kilic, T. (2017). How much of the labor in African agriculture is provided by women?. *Food policy*, 67, 52-63.

[41] Pretty, J. (1999). Can sustainable agriculture feed Africa? New evidence on progress, processes and impacts. *Environment, development and sustainability*, 1(3-4), 253-274.

[42] Ringler, C. & Sun, Y., 2010. Ethiopia Nile Basin Climate Change Adaptation Dataset, http://hdl.handle.net/1902.1/16273, Harvard Dataverse, V4

[43] Rosenbaum, P. R. & Rubin, D. B. (1983). The Central Role of the Propensity Score in Observational Studies for Causal Effects *Biometrika*, 70(1), 41-55

[44] Schlenker, W., & Lobell, D. B. (2010). Robust negative impacts of climate change on African agriculture. *Environmental Research Letters*, 5(1), 014010.

[45] Schultz, T. W. (1980). Nobel lecture: The economics of being poor. *The Journal of Political Economy*, 639-651. Chicago

[46] Shiferaw, B., & Holden, S. (1999). Soil erosion and smallholders' conservation decisions in the highlands of Ethiopia. *World development*, 27(4), 739-752.

[47] Singh, I., Squire, L., & Strauss, J. (1986). Agricultural household models: Extensions, applications, and policy. The World Bank.

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

[49] Taddese, G. (2001). Land degradation: a challenge to Ethiopia. *Environmental management*, 27(6), 815-824.

[50] Tesfaye, A., Brouwer, R., van der Zaag, P., & Negatu, W. (2016). Assessing the costs and benefits of improved land management practices in three watershed areas in Ethiopia. *International Soil and Water Conservation Research*, 4(1), 20-29.

[51] United Nations, Department of Economic and Social Affairs, Population Division (2015). World Population Prospects: The 2015 Revision, Methodology of the United Nations Population Estimates and Projections. ESA/P/WP.242.

[52] Zeng, D., Alwang, J., Norton, G. W., Shiferaw, B., Jaleta, M., & Yirga, C. (2014). Agricultural technology adoption and child nutrition: Improved maize varieties in rural Ethiopia. In *Selected Paper prepared for presentation at the Agricultural and Applied Economics Associations 2014 AAEA Annual Meeting.*

Chapter 5

The importance of a comprehensive drought index to estimate drought-induced cereal losses in India

Abstract

Drought events have critical impacts on agricultural production, yet there is little consensus on how these should be measured and defined. This has implications for research and policy, as drought is often defined purely based on rainfall. Recently, Babcock and Yu (2010) have developed an index that incorporates temperature. However, the authors focus uniquely on 'hot' droughts when temperature is included. We develop a flexible, rainfall-temperature drought index that captures all dry events, including a previously overlooked class of drought events: cold droughts. Our index is applied to a panel dataset of Indian districts over the period 1966-2009. Results suggest a statistically significant relationship between the index and agricultural production. Cold droughts are found to have consistent, negative marginal impacts that are comparable to those of hot droughts. Estimates of average yield losses due to hot drought are reduced by as much as 33% when cold droughts are omitted. The associated economic costs are even more severely underestimated, by up to 107%.

Keywords: Agriculture, Cereals, Climate, Drought, India, Rainfall, Temperature

JEL classification: Q10, Q19, Q54, Q56

5.1 Introduction

Extended periods of low rainfall and high temperature that reduce the availability of moisture relative to normal climate conditions broadly constitute drought events (Mishra and Singh, 2010). A number of low- and middle-income countries in the world, including those located in Sub-Saharan Africa and the Indian sub-continent, are particularly vulnerable to the impacts of such events. The human and economic costs of drought can be considerable. In India, the setting for our paper, Gadgil and Gadgil (2006) estimate that severe drought lowered annual GDP by around two to five percent between 1951 and 2003, while Pandey et al. (2007) show that drought was accompanied by a 12 to 33% increase in the poverty headcount ratio and a 25 to 60% decline in household income. The onset of drought in India has also been empirically linked to conflict, rural wages and human capital accumulation (Jayachandran, 2006; Sarson, 2015; Shah and Steinberg, forthcoming).

Against a backdrop of rising temperatures and drier conditions, drought is projected to become more common with critical implications for agricultural production (IPCC, 2012). How drought is defined plays a central role in policy-makers responses, not only in the agricultural sector but also in the water sector and in early-warning systems. Yet, in the academic and policy literatures there is little consensus on how drought might be measured and, hence, defined. Indeed, there is no universal definition of the conditions constituting a drought (Wilhite, 2000). A range of indices attempt to quantify the severity of a drought, ranging from simple rainfall measures to complex indices that account for rainfall, temperature and estimates of potential evapotranspiration (Mishra and Singh, 2010). Different criteria of what constitutes a drought therefore imply that a drought in one index may not constitute a drought in another. Thus, depending on the index used, there are classes of dry events which may simply be overlooked both in empirical analysis and by policymakers.

In this paper, we develop a simple rainfall-temperature index that allows for a flexible characterisation of drought events. It accounts for every dry event, in which cumulative (growing season) precipitation is below average, long-term cumulative (growing season) precipitation[1], while accounting for temperature. The novelty of our index is to include both the type of

[1] In our main specification, cumulative growing season precipitation is defined as the total rainfall between June-September. Long-term cumulative precipitation is defined as the average cumulative growing season precipitation between 1956-2009.

dry events captured by the Babcock and Yu (2010) index, i.e. characterised by above-average temperatures ('hot' drought), as well as ones characterised by below-average temperatures, which we term cold drought. Our index is applied to a panel dataset of Indian districts over the period 1966-2009 in order to estimate the marginal and total effects of drought on cereal productivity. These estimates are then used to simulate changes in yield and associated economic impacts. In a country where over two-thirds of total land area is vulnerable to drought (Ministry of Agriculture, 2009), and rain-fed agriculture covers approximately 60% of cropped area (Sharma, 2011), our analysis contributes to an important body of research on the impacts of drought on Indian agriculture (e.g. Pandey et al., 2007; Sarkar, 2011).

After motivating our analysis in the context of the relevant literature in Section 2, we present Indian weather data underlying hot and cold drought in Section 3. In Section 4, we propose an extension to an index originally developed by Yu and Babcock (2010). This extension allows for a more flexible characterization of drought events while retaining a key strength of their index, namely the inclusion of temperature. Applied to our panel dataset of Indian districts in Section 5, we find a statistically significant relationship between the index and agricultural production. We also find that cold droughts consistently display large negative marginal and total effects, comparable to those of hot droughts, and that omitting cold droughts leads to a large underestimation of total drought impact. Yield and economic losses are shown in Section 6 to be underestimated by up to 33% and 107%, respectively. Section 7 concludes.

5.2 Defining drought

Simple drought indices often rely solely on precipitation measures and are typically preferred by policy-makers including the Indian Meteorological Department (IMD) over more complex indices. The IMD recorded a drought event when seasonal rainfall was below 75% of its long-term average value (between 1950 and 2000), and a severe drought when rainfall was below 50% of this value. Simple metrics of precipitation deficiency, which have the advantage of being easily interpretable, are also used to evaluate drought impacts on agricultural production. For example, to estimate drought impact in the rice-growing regions of Asia, Pandey et al. (2007) define drought as moderate if rainfall is 70-80 percent of normal levels, and severe if rainfall is 70 percent below normal. Auffhammer et al. (2012) use a similar definition to study the effect of monsoon rainfall on rice yields in Indian states. The strength of these indices lies in their simplicity.

However, simple definitions of drought are problematic for our understanding of drought impacts for two reasons. First, they impose arbitrary thresholds in order to define drought, evaluating drought impacts only after a given level of precipitation, when the agronomic or empirical basis of such thresholds is unclear (Wilhite and Glantz, 1985). Second, variables in addition to precipitation, in particular temperature, help determine the physical severity of a drought[2]. Given temperature increases driven by climate change (Hatfield et al., 2011) a growing literature suggests critical turning points at which higher temperatures cease to have positive impacts on agricultural yield. Schlenker and Roberts (2009) find that higher temperatures in the US reduce county-level yields for corn (above 29°C), soy-beans (30°C), and cotton (32°C). Guiteras (2009) and Burgess et al. (2014) both show that, on average, daily temperatures above 34°C in India reduce agricultural productivity at the district scale. Lobell et al. (2012) identify the same threshold as harmful for Indian wheat yields.

High temperatures have particularly acute effects on crop growth during periods of low precipitation since the rate of evapotranspiration, i.e. the combined process of water evaporated from land surfaces and plants, increases as temperatures rise (Prasad et al., 2008; Lobell and Gourdji, 2012). In general, this increases a plant's demand for water at a time when water availability is already low due to deficient precipitation. Recent research has documented that droughts in a range of settings have increased in severity as mean temperatures have risen. Higher temperatures, rather than the increased intensity of low rainfall events, have been held responsible for these drying trends (Vicente-Serrano et al., 2014; Diffenbaugh et al., 2015). As such, not considering the effect of temperature on the severity of a drought event could underestimate drought impact, in turn giving misleading information about the likelihood of future production losses driven by climate change.

More complex indices tend to rely on data that are not readily available in most economic datasets, e.g. for soil moisture levels and estimates of potential evapotranspiration, which can depend on factors such as wind, radiation and humidity, thus limiting their applicability in empirical analysis of drought impacts. In an attempt to bridge the gap between simple

[2]Such variables include, for example, the access to irrigation.

and complex indices, Yu and Babcock (2010) propose a drought index that neatly captures the interaction between temperature and precipitation. Applied to the study of resilience of soybean and corn yields in the US, it takes a non-zero value only for years of below-average precipitation and above-average values of heat temperature (cooling degree-days). The authors find that soybeans and corn have become increasingly drought-tolerant over time.

This index has since been applied in a number of other settings, for example, to the assessment of drought impact on soybean in Missouri (Purcell and Caine, 2013). Of particular relevance is research by Birthal et al. (2015), who use the index to study the resilience of rice yields to drought in India. Their results indicate that rice yields have become more resilient to drought over time. While this approach has the advantage of being a relatively simple way to account for both temperature and precipitation, the index restricts the definition of drought to events characterised by low rainfall accompanied by higher-than-average temperatures. It does not consider events characterised by below-mean rainfall as well as below-mean temperature. Such cold droughts are common to many settings, although their impacts on agricultural production remain unknown, due to either being omitted altogether (as in Birthal et al. 2015) or joined with hot droughts in arbitrarily-defined rainfall indices. This is an important gap in the literature that our paper aims to fill.

We argue that cold droughts should not be omitted *a priori* for two reasons. First, a large number of potentially destructive droughts are not considered, which can lead to a serious underestimation of the total impact of drought. Second, the classification of these events as non-droughts could lead to biased estimates of drought impact. Thus, if cold droughts have a significant negative impact on productivity, Yu and Babcock's (2010) index is likely to underestimate drought impacts because events that we might define as cold drought are included in their control group.

5.3 Drought in India

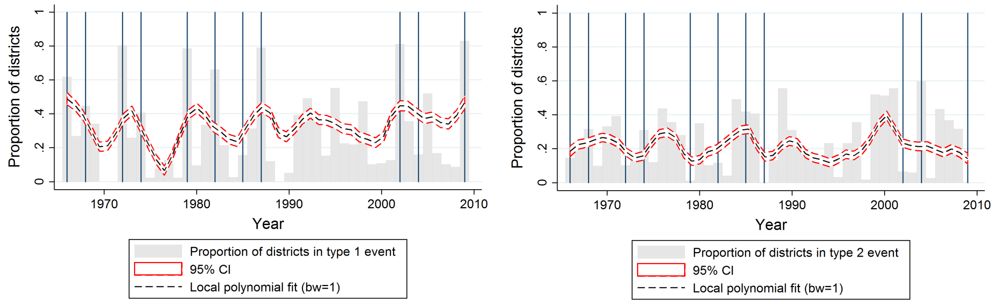
According to the definition used by the IMD (at least until 2016), 13 All-India drought years have been recorded since the beginning of the Green Revolution in 1966 (Birthal et al., 2015). Four of these occurred between 2000 and 2012. A drought year was recorded when the total area affected by a moderate or severe drought covered 20-40% of the total land area of the country and seasonal rainfall deficiency during the monsoon season exceeded 10%. When more than 40% of the total land area was affected by drought, this was known as an "All India Severe Drought Year".

Weather data on daily rainfall and daily average temperatures at the district level are sourced from the IMD to create Figures 5.1 and 5.2[3]. Panel (a) of Figure 5.1 shows the proportion of districts in every given year limited to events characterized by both below-average rainfall and above-average temperature, i.e. the events considered by Birthal et al. (2015) using Yu and Babcocks (2010) index. The vertical blue lines indicate the years defined by Indias government as 'All-India' droughts. Panel (b) of Figure 5.1 shows the proportion of districts in years characterised by below-average rainfall and below-average temperature; a large proportion of districts are clearly affected by this type of drought event. Figure 5.2 shows why the omission of these events is problematic. For each year, we estimate the number of districts affected by hot droughts net of the number of those affected by cold droughts, with a positive number (in red) denoting a year in which the former exceeds the latter. A negative number (in blue) indicates a year in which the latter exceeds the former. Overall, hot droughts are slightly more prevalent than cold droughts (roughly a split of 55% hot and 45% cold drought). In the 1990s, most of the drought-affected districts were affected by hot droughts. Since 1999, the number of cold droughts has increased, with the number of districts affected by cold droughts outnumbering districts affected by hot droughts in seven out of 11 years[4].

[3]The rainfall data are available in gridded format at a resolution of 0.25°x 0.25° (Pai et al., 2014). Gridded temperature data are at a resolution of 1°x1° (Srivastava et al., 2009). District-level weather data are then obtained by taking a weighted average of gridded weather observations from grid cells that fall within a district's boundary based on the proportion of the grid cell that falls in each district.

[4]This pattern, however, is slightly less pronounced if we look at an alternative growing season (May-December) (Figures 5A.1 and 5A.2 in the Appendix).

Figure 5.1: Proportion of drought-affected districts (by type) (June-September only)

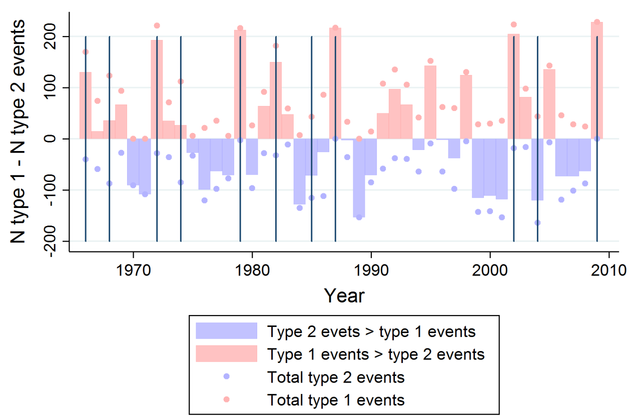


(a)Type 1 events ("hot droughts") (below-average rain & above-average temperature)

(b)Type 2 events ("cold droughts") (below-average rain & below-average temperature)

Notes: Type 1 events denote events where rainfall was below-average and temperature was above-average. Conversely, Type 2 events refer to events where both rainfall and temperature were below-average. The rainfall average variable is calculated as the district mean cumulative rainfall from June-September from 1956-2009. The average temperature variable is calculated as the average degree days above the mean season temperature from June-September. The solid vertical lines represent the years considered by the Indian Government as All-India drought years. Source: Authors' own calculations

Figure 5.2: Type 1 droughts in excess of Type 2 droughts (June-September only)



Notes: The scatter points highlight the total number of droughts (by type) in a given year. In the case of Type 1 droughts (red scatter points) these can be interpreted directly (i.e. 200 means that 200 districts were affects by a Type 1 drought). However, in the case of the Type 2 droughts, these should be interpreted as the negative of the number (i.e. if the observed value is -100, this means there were 100 districts affected by Type 2 droughts). Bar graphs show the number of affected districts affected by Type 1 droughts in excess of the number affected by Type 2 droughts. As a result, a value of 50 would mean that there were 50 more districts affected by a Type 1 drought than affected by a Type 2 drought in a given year. The converse applies to a negative number, which highlights a higher number of districts affected by cold droughts in a given year. The solid vertical lines represent the years considered by the Indian Government as All-India drought years. Source: Authors own calculations

5.4 Introducing a new drought index

In this section, we build on Yu and Babcocks (2010) drought index, incorporating both rainfall and temperature. Their index is based on the following:

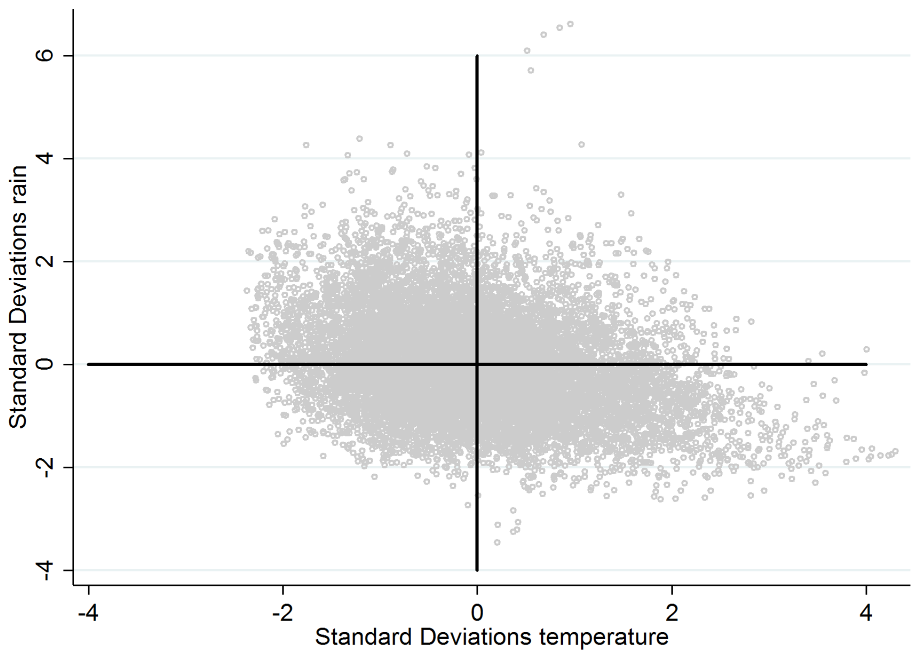
$$DI_{i,t} = [-max(0, CLDD^{stand}_{i,t})] * [min(0, TPCP^{stand}_{i,t})] \tag{5.1}$$

where: DI denotes the drought index, for a given unit of observation, i, in year t; $CLDD^{stand}_{i,t}$ is standardized Cooling Degree Days (above 65°F, or 18.33°C); and,$TPCP^{stand}_{i,t}$ is standardized total monthly precipitation between the months of June and August.

This index gives a value of zero to drought whenever either the temperature is below average or the rainfall is above average. As such, a drought occurs in a year when temperature is uncommonly high and precipitation is low, relative to the long-term average of these variables. The strength of this index lies in its capacity to capture the potential of high temperatures to exacerbate the effects of low rainfall on crop production, in a simple way.

One weakness of the index described in 5.1 is that it defines as a drought only those years when an area suffers both low rainfall and high temperatures. Omitted are years when rainfall is low but temperatures are not particularly high. Defining drought events by low rainfall and high temperature restricts the measure of drought to the lower-right quadrant (which we denote as quadrant 1) of Figure 5.3 . Events in the lower-left quadrant (which we denote as quadrant 2), where both the precipitation and temperatures are below-average, are not considered droughts.

Figure 5.3: Rainfall-temperature quadrants



Notes: Long-term average rainfall is calculated as the average cumulative rainfall for the June-September period for the 1956-2009 period. Average temperature is defined as the number of degree-days above the mean daily June-September temperature for the 1956-2009 period.

We consider a wider set of drought events by defining six variables. First, using weather data from the IMD, we calculate district-specific average long-term cumulative rainfall, $LTAR_i$, for the growing season (June-September) over the period 1956-2009[5]. This variable is standardized by estimating $ZTR_{it} = \dfrac{TR_{it} - LTAR_i}{sdTR_i}$, where TR_{it} is total cumulative rainfall over the growing season for a given year and $sdTR_i$ is the standard deviation of TR_{it}. Analogously, we calculate the district-specific average cumulative growing season temperature, $LTAHDD_i$, for the growing season (June-September) as the average cumulative number of daily degree days above the mean daily growing season temperature over the period 1956-2009[6]. Similar to

[5]There are two main reasons driving our choice of growing season. First, the majority of India's cereal production is cultivated in the kharif season, between June and September. Second, according to Jain and Kumar (2012), the majority of total yearly rainfall (approximately 80%) occurs between June and September. Authors such as Prasana (2014) also highlight that, while there is a strong and positive response to kharif production and June-September rainfall, the same is not necessarily true for rabi production and post-monsoon rainfall (October-December). This partly relates to the fact that rabi crops rely on available moisture from the June-September rains. The sensitivity of our results is tested for two alternative growing seasons in Section 5, namely using cumulative rainfall for the May-December period and using cumulative annual rainfall.

[6]The growing season daily degree days are calculated as follows. First, we obtain the average growing season temperature. Second, for each day we subtract the average temperature from the observed temperature and obtain the number of degrees above the average temperature for each day. Finally, we sum all the positive

rainfall, this variable is standardized by estimating $ZHDD_{it} = \dfrac{HDD_{it} - LTAHDD_i}{sdHDD_i}$, where HDD_{it} is total cumulative daily degree days over the growing season for a given year and $sdHDD_i$ is the standard deviation of HDD_{it}.

Let $MTR_{it} = -TR_{it}$, i.e. the negative of total cumulative rainfall. We then obtain the normalized version of this variable, NTR_{it}, by estimating $NTR_{it} = \dfrac{MTR_{it} - MTR_i^{min}}{MTR_i^{max} - MTR_i^{min}}$, where MTR_i^{min} denotes the minimum observed value for district i (i.e. the maximum rainfall observed), and MTR_i^{max} denotes its maximum observed value (i.e. lowest rainfall). Normalizing the negative of rainfall, rather than rainfall directly, allows us to generate a variable bounded between 0 and 1, with higher values signalling a more severe precipitation deficiency. Similarly, for normalizing degree days we estimate $NHDD_{it} = \dfrac{HDD_{it} - HDD_i^{min}}{HDD_i^{max} - HDD_i^{min}}$, where HDD_i^{min} denotes the minimum observed value for district i (i.e. the minimum number of degree-days observed), and HDD_i^{max} denotes its maximum observed value (i.e. highest number of degree days observed).

A multiplicative relationship is generated between the two normalized variables, which we use to define three different drought indices. First, hot droughts can be classified as $D1_{it}$, corresponding to the classification of Yu and Babcock (2010), where rainfall is below normal and temperature above normal. Second, $D2_{it}$ corresponds to low rainfall in the absence of abnormally high temperatures. Third, we combine $D1_{it}$ and $D2_{it}$ to get $D12_{it}$, thus accounting for both hot and cold droughts. Formally, we have:

$$Drought = \begin{cases} D1_{it} = NTR_{it} * NHDD_{it} \text{ if } ZTR_{it} < 0 \textit{ and } ZHDD_{it} > 0; \;\; 0 \text{ otherwise} \\ D2_{it} = NTR_{it} * NHDD_{it} \text{ if } ZTR_{it} < 0 \textit{ and } ZHDD_{it} < 0; \;\; 0 \text{ otherwise} \\ D12_{it} = NTR_{it} * NHDD_{it} \text{ if } ZTR_{it} < 0; \;\; 0 \text{ otherwise} \end{cases} \tag{5.2}$$

As such, $D1_{it}$ can be interpreted as a normalized version of Yu and Babcock's (2010) index. It captures all events in the lower right quadrant (quadrant 1) of Figure 5.3, taking a strictly positive value for all events characterized by below-average precipitation and above-average temperatures. The second index, $D2_{it}$, only takes non-zero values for events with below-

temperature deviations for each day of the growing season and obtain the cumulative daily-degree days.

average rainfall and below-average temperature, the category Yu and Babcock (2010) omit. Constructing these two indices separately allows us to test their respective statistical significance in the yield regressions in Section 5. Finally, a third index, $D12_{it}$, simply combines $D1_{it}$ and $D2_{it}$ and hence, captures all the events in the lower half of Figure 5.3. More detail on how the indices are constructed is presented in Appendix B.

Our indices are increasing in temperature but decreasing in precipitation since both higher temperatures and lower precipitation are expected to contribute to drought severity. A maximum value of 1 is obtained when drought is most severe, and is only possible for the restricted set of drought events considered by Yu and Babcock. The similarity of their index to our own is illustrated in Table 5.1, which shows the correlation coefficients and the spearman correlation coefficient[7]. As expected, our index for quadrant 1, D1, is highly correlated with Yu-Babcock, displaying a correlation coefficient of 0.776 and a spearman correlation coefficient in excess of 0.99. Our second index, on the other hand, has a negative correlation coefficient. Since Yu-Babcock is invariant with a value of zero for these events, this result is also as anticipated.

Table 5.1: Correlation coefficients and Spearman correlation coefficients

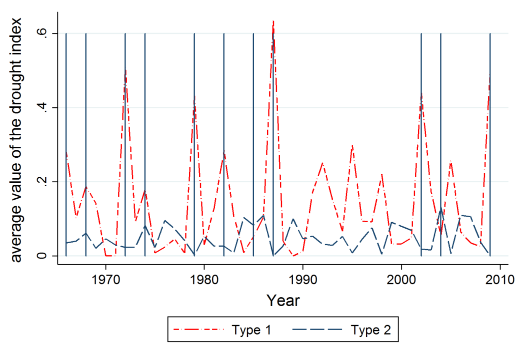
Correlation coefficients				
	Babcock-Yu	DI (q1)	DI (q2)	DI (q1 and q2)
Babcock-Yu	1.000			
DI (q1)	0.776	1.000		
DI (q2)	-0.181	-0.303	1.000	
DI (q12)	0.736	0.919	0.097	1.000
Spearman correlation coefficients				
	Babcock-Yu	DI (q1)	DI (q2)	DI (q1 and q2)
Babcock-Yu	1.000			
DI (q1)	0.992	1.000		
DI (q2)	-0.359	-0.359	1.000	
DI (q12)	0.821	0.830	0.217	1.000

Figure 5.4 shows how our indices change over time, for all districts (panels (a) and (b)) and for drought-affected districts only (panels (c) and (d)). Hot (D1) and cold (D2) droughts

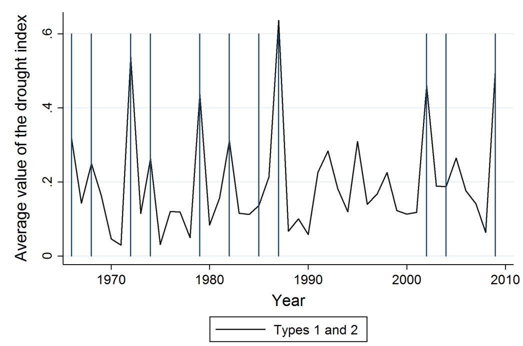
[7]Note that, instead of cooling days, we use hot degree days based on the mean temperature of the district for the months between June and September over the 1956-2009 period.

are denoted Types 1 and 2, respectively. There are clear spikes in the values of the index for a number of years considered All-India drought years. In recent years, 2002 and 2009 are associated with the largest deviations in rainfall; spikes correspond to these two years. Similarly, 1972, 1979, 1987 are also considered years with particularly high deviations and our index rises in these years. Throughout the 1990s, however, it is striking that, despite relatively modest deviations of rainfall from trend, our index still records high values. On average, the negative deviations from long-term average rainfall were smaller throughout the 1990s. A possible explanation for this could be the fact that, as highlighted by Pai et al. (2012), overall, land surface air temperatures have increased over time. This pattern was particularly pronounced in the 1990s and 2000s.

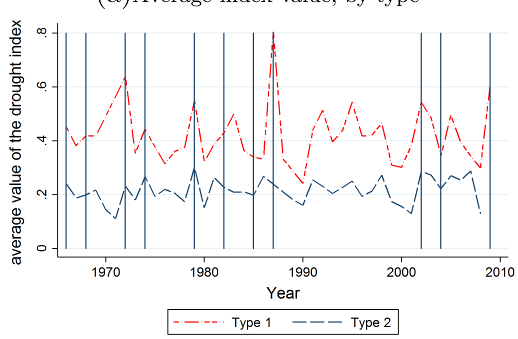
Figure 5.4: Average drought index value



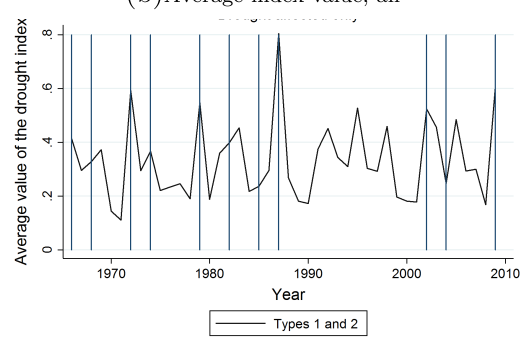
(a) Average index value, by type



(b) Average index value, all



(c) Average index value, by type (affected only)



(d) Average index value, all (affected only)

5.5 Impact of drought on cereal productivity

5.5.1 Data and Methodology

To investigate drought impacts on aggregate cereal productivity at the district level, we obtain agricultural data from the ICRISAT Meso-level Database[8]. For the period 1966-2009, the dataset contains detailed agricultural and socioeconomic information (ICRISAT, 2012). For most if not all districts, data are available for annual crop production and area under crop production for a range of crops. We create a balanced panel, which implies that, out of the 311 districts available in the dataset, only 275 districts are used in our empirical analysis due to missing weather and/or production data. Six cereals are considered, namely rice, wheat, maize, barley, sorghum, and millet[9]. Yields for each are estimated along with a simple cereal yield variable, obtained by dividing total cereal production by total cereal area. Table 5.2 summarises the variables used in our empirical analysis.

[8] Since 1966, a number of districts have split into smaller districts. To maintain spatial consistency over time district, splits are dealt with by returning split districts to their parent districts as of 1966.

[9] For millet we add data on quantities of pearl millet and finger millet to create an aggregate quantity of millet.

Table 5.2: Summary Statistics of observations in the sample

Variables	**N**	**Mean**	**S.D**	**Min**	**Max**
Cereal yield (t/ha)	12100	1.463	0.787	0.006	4.775
Cereal Area (1,000,000 ha)	12100	0.332	0.195	0.001	1.334
Barley yield (t/ha)	5842	1.383	0.688	0.048	5.400
Cereal area under barley production (%)	12031	0.015	0.036	0.000	0.320
Maize yield (t/ha)	10621	1.431	0.929	0.003	9.739
Cereal area under maize production (%)	12099	0.065	0.113	0.000	0.838
Millet yield (t/ha)	9852	0.798	0.439	0.000	4.000
Cereal area under millet production (%)	12100	0.131	0.220	0.000	1.000
Rice yield (t/ha)	11398	1.492	0.853	0.009	5.542
Cereal area under rice production (%)	12100	0.401	0.357	0.000	1.000
Sorghum yield (t/ha)	9694	0.774	0.434	0.001	9.836
Cereal area under sorghum production (%)	12066	0.148	0.225	0.000	0.929
Wheat yield (t/ha)	10275	1.643	0.878	0.046	6.324
Cereal area under wheat production (%)	12093	0.240	0.246	0.000	0.972
Proportion of net irrigated area (%)	12095	0.355	0.270	0.000	1.467
Rural population density (by gross cereal area)	11787	3.566	2.142	0.428	17.907
Fertiliser (t/1,000 ha)	11889	60.571	61.406	0.000	614.493
Cumulative rainfall (mm) (June-September)	12100	863.837	529.348	13.125	5313.428
Hot Degree-Days (HDD, June-September)	12100	94.422	47.204	2.697	278.413
Babcock-Yu index, June-September	12100	0.270	0.752	0.000	7.998
Drought index (quadrant 1)	12100	0.146	0.245	0.000	1.000
Drought index (quadrant 2)	12100	0.049	0.097	0.000	0.544
Drought index (quadrants 1 and 2)	12100	0.196	0.237	0.000	1.000

Notes: Rural population density is calculated by dividing total rural population by gross cropped area. Our hot degree-days measure is calculated based on average daily district temperature in the months of June-September for the period 1956-2009.

To model the relationship between yield and our drought index, we estimate the following fixed-effects model:

$$ln(y_{itc}) = \alpha_i + \gamma_t + \delta_{i1} * t + \delta_{i2} * t^2 + \beta_q DI_{itq} + \epsilon_{it} \tag{5.3}$$

where for district i in year t: $ln(y_{itc})$ denotes the natural logarithm of cereal yield (or crop c); α_i and γ_t represent the district and year fixed effects, respectively; δ_{i1} and δ_{i2} are the coefficients on the district-specific quadratic trend. The coefficient associated with a type q (i.e. Type 1 - hot or Type 2 - cold or Type 12 - both types) drought index, which captures the marginal impact of a type q drought, is denoted β_q. Finally, ϵ_{it} represents the error term. Consistent with Yu and Babcock (2010), we do not include controls in our main specifications. This is also the norm in the broader weather and climate literature. Nevertheless, we test the sensitivity of our results to the inclusion of controls (see next section and Appendix tables).

5.5.2 Regression results

We run a regression of the natural logarithm of yield on a set of district-specific quadratic trends and the drought indices. Specifically, for each specification we estimate three different regressions[10]. First, we include only Type 1 drought events. Second, we estimate separate coefficients for Type 1 and Type 2 drought events. Finally, we run a regression where we only include the drought index that combines Type 1 and Type 2 events. The results for the full sample can be seen in Table 5.3. The results by crop are in Tables 5.4-5.5 (without dummies) and 5.6-5.7 (with dummies).

Table 5.3 highlights three main points. First, both types of drought have significant and negative effects when considered separately, as shown in columns 2 and 5. In practice, this means that Type 2 events, i.e. those omitted by Yu and Babcock (2010) and Birthal et al. (2015), have large and statistically significant, negative impacts on yield.

[10]For each model (main and robustness) we run the regression for both the full sample and by crop. For both the full sample and for each crop we then run three regressions.

Table 5.3: Full sample results

	No dummies			Dummies		
Variables	1	2	3	4	5	6
Dummy (Type 1)				0.041**	0.019	
				(0.018)	(0.018)	
Drought index (Type 1)	-0.192***	-0.238***		-0.267***	-0.271***	
	(0.019)	(0.021)		(0.041)	(0.042)	
Dummy (Type 2)					0.022**	
					(0.011)	
Drought index (Type 2)		-0.391***			-0.471***	
		(0.036)			(0.052)	
Dummy (2 types)						-0.011
						(0.009)
Drought index (2 types)			-0.255***			-0.232***
			(0.021)			(0.028)
Constant	-0.361***	-0.326***	-0.313***	-0.359***	-0.323***	-0.316***
	(0.022)	(0.022)	(0.022)	(0.022)	(0.022)	(0.023)
Time trends	✓	✓	✓	✓	✓	✓
District fixed effects	✓	✓	✓	✓	✓	✓
Year fixed effects	✓	✓	✓	✓	✓	✓
Controls						
Number of observations	12100	12100	12100	12100	12100	12100
Number of districts	275	275	275	275	275	275
R-squared a	0.705	0.712	0.711	0.705	0.713	0.711
R-squared w	0.719	0.726	0.725	0.72	0.727	0.725

Notes: Values in parentheses denote clustered standard errors at the district level. *, ** and *** denote statistical significance at the 10%, 5% and 1% level, respectively. District trends denote a quadratic district-specific trend.

Second, by comparing the specifications where we omit Type 2 events (columns 1 and 4) with those where this type of event is included (columns 2 and 5), we note that the estimated marginal coefficient of Type 1 droughts is smaller in the former. When we include a dummy variable, the difference in magnitude is negligible. But when all dummy variables are excluded, the coefficient of Type 1 events is substantially smaller - and outside the 95% confidence interval of the estimated coefficient - when Type 2 events are also included (for a graphical representation, see Figure 5A.3 in the Appendix). Thus, a failure to account for Type 2 events can lead to an underestimation of the marginal impacts of drought, with this underestimation being more severe when a dummy is not included.

Third, we find that cold droughts have a larger marginal, but lower total, effect on agricultural production. Although both excess heat and reduced moisture have negative impacts on production, reduced precipitation carries greater weight in the cold drought index than in the hot drought index since values of temperature are, by definition, higher in the latter than in the former. As a result, yields are likely to respond (more) negatively to changes in the cold drought index than in the hot drought index. A value of 0.5 in our cold drought index represents approximately the same precipitation deficiency as a value of 1 in our hot drought index, which could help explain larger marginal impacts. This last result (larger marginal effects in the case of cold droughts), however, is not robust in a number of alternative specifications.

The results by crop (Tables 5.4-5.7) corroborate the patterns found across the whole sample using the cereal index. The estimated coefficients for Type 2 events are consistently large, negative and significant for all crops except for maize (when dummies are excluded). This provides further evidence that such events have a large negative impact on production and, hence, should not be excluded from analyses of drought impact when considering any of the most important cereal crops grown in India. Similar to our findings for the whole sample, in most cases (maize again being the exception) when a dummy is not included to account for the intercept shift, the omission of Type 2 events leads to a smaller, estimated coefficient of Type 1 droughts. This effect is especially large in the case of rice, the crop analysed by Birthal et al. (2015), thus implying that they may have underestimated the impact of drought on rice. We estimate the potential scale of underestimation, in terms of yield and its economic value, below.

Table 5.4: Results by crop - rice, wheat and maize (no dummies)

	Rice			Wheat			Maize		
Variables	1	2	3	4	5	6	7	8	9
Drought index (Type 1)	-0.249***	-0.303***		-0.128***	-0.161***		-0.138***	-0.118***	
	(0.020)	(0.023)		(0.016)	(0.017)		(0.026)	(0.029)	
Drought index (Type 2)		-0.469***			-0.263***			0.156***	
		(0.041)			(0.032)			(0.056)	
Drought index (2 types)			-0.323***			-0.174***			-0.084***
			(0.024)			(0.017)			(0.029)
Constant	-0.234***	-0.197***	-0.180***	-0.339***	-0.311***	-0.299***	-0.201***	-0.218***	-0.247***
	(0.034)	(0.034)	(0.034)	(0.027)	(0.028)	(0.029)	(0.050)	(0.050)	(0.050)
Time trends	✓	✓	✓	✓	✓	✓	✓	✓	✓
District fixed effects	✓	✓	✓	✓	✓	✓	✓	✓	✓
Year fixed effects	✓	✓	✓	✓	✓	✓	✓	✓	✓
Controls									
Number of observations N	10560	10560	10560	8756	8756	8756	7656	7656	7656
Number of districts	240	240	240	199	199	199	174	174	174
R-squared a	0.532	0.542	0.541	0.713	0.716	0.716	0.398	0.398	0.396
R-squared w	0.555	0.565	0.563	0.727	0.731	0.73	0.428	0.429	0.426

Notes: Values in parentheses denote clustered standard errors at the district level. *, ** and *** denote statistical significance at the 10%, 5% and 1% level, respectively. District trends denote a quadratic district-specific trend.

Table 5.5: Results by crop - millet, sorghum and barley (no dummies)

	Millet			Sorghum			Barley		
Variables	1	2	3	4	5	6	7	8	9
Drought index (Type 1)	-0.240***	-0.287***		-0.159***	-0.188***		-0.013	-0.032	
	(0.039)	(0.044)		(0.031)	(0.035)		(0.027)	(0.029)	
Drought index (Type 2)		-0.381***			-0.239***			-0.152***	
		(0.078)			(0.066)			(0.045)	
Drought index (2 types)			-0.297***			-0.195***			-0.051*
			(0.045)			(0.036)			(0.027)
Constant	-0.724***	-0.692***	-0.684***	-0.728***	-0.710***	-0.703***	-0.349***	-0.328***	-0.315***
	(0.041)	(0.041)	(0.041)	(0.051)	(0.050)	(0.048)	(0.040)	(0.040)	(0.039)
Time trends	✓	✓	✓	✓	✓	✓	✓	✓	✓
District fixed effects	✓	✓	✓	✓	✓	✓	✓	✓	✓
Year fixed effects	✓	✓	✓	✓	✓	✓	✓	✓	✓
Controls									
Number of observations	7172	7172	7172	6908	6908	6908	3432	3432	3432
Number of districts	163	163	163	157	157	157	78	78	78
R-squared a	0.429	0.434	0.434	0.335	0.337	0.337	0.767	0.768	0.767
R-squared w	0.459	0.463	0.463	0.369	0.371	0.371	0.78	0.781	0.78

Notes: Values in parentheses denote clustered standard errors at the district level. *, ** and *** denote statistical significance at the 10%, 5% and 1% level, respectively. District trends denote a quadratic district-specific trend.

Table 5.6: Results by crop - rice, wheat and maize (with dummies)

	Rice			Wheat			Maize		
Variables	1	2	3	4	5	6	7	8	9
Dummy (Type 1)	0.047	0.014		0.01	-0.006		0.157***	0.185***	
	(0.033)	(0.033)		(0.017)	(0.017)		(0.032)	(0.031)	
Drought index (Type 1)	-0.336***	-0.330***		-0.145***	-0.150***		-0.424***	-0.443***	
	(0.062)	(0.061)		(0.039)	(0.039)		(0.068)	(0.068)	
Dummy (Type 2)		-0.023*			0.01			0.169***	
		(0.013)			(0.011)			(0.027)	
Drought index (Type 2)		-0.380***			-0.300***			-0.439***	
		(0.072)			(0.057)			(0.107)	
Dummy (2 types)			-0.033***			-0.016**			0.169***
			(0.010)			(0.008)			(0.018)
Drought index (2 types)			-0.257***			-0.143***			-0.417***
			(0.029)			(0.024)			(0.046)
Constant	-0.232***	-0.199***	-0.191***	-0.339***	-0.309***	-0.305***	-0.198***	-0.187***	-0.184***
	(0.034)	(0.035)	(0.035)	(0.027)	(0.028)	(0.029)	(0.050)	(0.051)	(0.051)
Time trends	✓	✓	✓	✓	✓	✓	✓	✓	✓
District fixed effects	✓	✓	✓	✓	✓	✓	✓	✓	✓
Year fixed effects	✓	✓	✓	✓	✓	✓	✓	✓	✓
Controls									
Number of observations	10560	10560	10560	8756	8756	8756	7656	7656	7656
Number of districts	240	240	240	199	199	199	174	174	174
R-squared a	0.532	0.542	0.541	0.713	0.716	0.716	0.4	0.406	0.406
R-squared w	0.556	0.565	0.564	0.727	0.731	0.73	0.431	0.436	0.436

Notes: Values in parentheses denote clustered standard errors at the district level. *, ** and *** denote statistical significance at the 10%, 5% and 1% level, respectively. District trends denote a quadratic district-specific trend.

Table 5.7: Results by crop - millet, sorghum and barley (with dummies)

	Millet			Sorghum			Barley		
Variables	1	2	3	4	5	6	7	8	9
Dummy (Type 1)	0.075**	0.062*		0.099**	0.093**		-0.005	-0.017	
	(0.031)	(0.031)		(0.043)	(0.045)		(0.024)	(0.024)	
Drought index (Type 1)	-0.378***	-0.392***		-0.340***	-0.351***		-0.003	0	
	(0.078)	(0.079)		(0.088)	(0.088)		(0.058)	(0.058)	
Dummy (Type 2)		0.087***			0.097***			0.004	
		(0.020)			(0.021)			(0.018)	
Drought index (Type 2)		-0.699***			-0.571***			-0.174**	
		(0.117)			(0.096)			(0.087)	
Dummy (2 types)			0.038**			0.059***			-0.022
			(0.017)			(0.022)			(0.014)
Drought index (2 types)			-0.371***			-0.310***			-0.004
			(0.058)			(0.050)			(0.046)
Constant	-0.727***	-0.690***	-0.677***	-0.726***	-0.692***	-0.679***	-0.350***	-0.327***	-0.324***
	(0.041)	(0.040)	(0.041)	(0.052)	(0.050)	(0.051)	(0.040)	(0.041)	(0.040)
Time trends	✓	✓	✓	✓	✓	✓	✓	✓	✓
District fixed effects	✓	✓	✓	✓	✓	✓	✓	✓	✓
Year fixed effects	✓	✓	✓	✓	✓	✓	✓	✓	✓
Controls									
Number of observations	7172	7172	7172	6908	6908	6908	3432	3432	3432
Number of districts	163	163	163	157	157	157	78	78	78
R-squared a	0.43	0.435	0.434	0.336	0.339	0.338	0.767	0.767	0.767
R-squared w	0.459	0.465	0.463	0.37	0.373	0.372	0.78	0.781	0.781

Notes: Values in parentheses denote clustered standard errors at the district level. *, ** and *** denote statistical significance at the 10%, 5% and 1% level, respectively. District trends denote a quadratic district-specific trend.

We perform a number of sensitivity checks on our results, including: (i) standard errors clustered at the state level; (ii) two alternative growing seasons (May-December and annual)[11]; (iii) an alternative specification for hot degree-days (30 degrees rather than the long-term district average temperature over the growing season)[12]; (iv) controls[13]; (v) alternative func-

[11] We use the May-December growing season in order to allow for the fact that, in some states, there may be substantial amounts of rain outside of the June-September period. The yearly growing season is used as the rainfall in the later months of the year may also be useful in explaining the rabi production. However, according to authors such as Prasana (2014), the dependence of Rabi production on post-monsoon rainfall is not as strong, which is one of the reasons why we opt for the June-September period as the growing season.

[12] We test for a different specification on temperature for two main reasons. According to research on the effects of temperature on crop yield in India, crop yields seem to start decreasing at different levels, but usually after 30 degrees Celsius. In our paper, however, we use mean temperature, which in some cases is likely to be under 30. As a result, it could be argued that an increase in temperature should not necessarily result in a decrease in production. However, there are two reasons why this is not the main choice in our paper. First, using 30 as a cut-off period means that the drought index will have a value of 0 in many drought events (when rainfall is very low) as we are using an absolute cut-off level, which makes our index less comprehensive. Secondly, districts with low temperature may also have adopted crops whose negative impacts start at lower levels of temperature and, as such, it is possible that, in some cases, we may observe negative impacts of temperature at levels below 30 degrees.

[13] This robustness check is carried out to ensure that our results are robust, even after including variables

tional forms of the index (including the square of the index)[14] ; and, (vi) an additive index instead of a multiplicative index[15]. For the full sample, the coefficients from each of these specifications are summarized in Tables 5A.1-5A.2. While different specifications unsurprisingly generate different coefficients, our overarching conclusions are quite robust, especially in the case of aggregate cereal production. In Figure 5A.3, we present a graphical summary of the coefficient values and their confidence intervals for each specification using the full sample. Tables 5A.3- 5A.14 summarize all the robustness checks by crop. Our results are robust, especially to alternative index specifications (Tables 5A.4, 5A.6, 5A.8, 5A.10, 5A.12, 5A.14).

5.6 Estimating yield and economic losses

In a back-of-the-envelope attempt to gauge first how important both types of droughts are in the Indian context and, second, how serious the omission of Type 2 droughts is for estimating Type 1 drought impacts, we run simple simulations using our estimated regressions. This allows us to generate predictions of yields with and without droughts, more specifically, to estimate the: (i) average yield loss for an affected district over the sample period; (ii) average total production loss for an affected district over the sample period; (iii) average total value of production for an affected district; (iv) average unweighted yearly total production loss in our sample of Indian districts; and, (v) the average yearly total cost across sampled districts. A summary of estimates is presented in Tables 5.8 and 5.9, including crop-specific results[16]. Details of how we generated these estimates can be found in Appendix C.

From tables 5.8 and 5.9, we note the following. Despite a higher estimated coefficient, total yield and economic losses from cold droughts are smaller than those from hot droughts. This is due to the index values for cold droughts being substantially lower (approximately half) for affected districts. For our aggregate cereal measure, we estimate the average yield loss per

which are also likely to affect production, such as irrigation and use of modern inputs.

[14]The impacts of drought may not be linear. As a result, we include a squared term to account for this.

[15]While Babcock and Yu (2011) advocate for a multiplicative relationship, an additive relationship may also be plausible. Moreover, the additive relationship is likely to be preferable in a number of extreme cases. For example, suppose that there is a year where rainfall is close to 0 (undoubtedly a bad year) but the temperatures have also been low. Our index would have a low value. However, in such a case, the multiplicative index is not capturing the fact that this has been a bad year, whereas the additive index is. However, the main rationale for using the multiplicative index is that we are capturing the interaction between rainfall and temperature.

[16]For the crop-specific results, these were obtained using the crop-specific regressions.

Table 5.8: Cost estimates - No dummies

	Main				
	Type 1 (only)	Type 1 (sep)	Type 2 (sep)	2 types (sep)	2 types (Together)
			Full sample		
Av. yield loss (district) (t/ha)	0.12	0.16	0.11	0.14	0.12
Av. production loss (district) (1,000t)	40.49	50.95	39.11	45.57	41.28
Av. production cost (district) (mil usd)	9.78	12.30	9.44	11.00	9.97
Av. yearly total production loss (1,000 t)	3,332.87	4,229.59	2,689.97	6,919.56	6,270.95
Av. yearly total cost (mil usd)	804.79	1,021.32	649.55	1,670.87	1,514.25
			Rice		
Av. yield loss (district) (t/ha)	0.16	0.20	0.14	0.17	0.16
Av. production loss (district) (1,000t)	24.56	30.55	21.41	26.38	24.11
Av. production cost (district) (mil usd)	7.36	9.15	6.41	7.90	7.22
Av. yearly total production loss (1,000 t)	1,745.70	2,177.00	1,283.33	3,460.33	3,185.97
Av. yearly total cost (mil usd)	522.79	651.95	384.32	1,036.28	954.11
			Wheat		
Av. yield loss (district) (t/ha)	0.10	0.13	0.09	0.11	0.10
Av. production loss (district) (1,000t)	10.99	14.04	11.32	12.77	11.56
Av. production cost (district) (mil usd)	2.46	3.14	2.53	2.86	2.58
Av. yearly total production loss (1,000 t)	655.71	822.57	580.93	1,403.51	1,272.08
Av. yearly total cost (mil usd)	146.62	183.93	129.90	313.83	284.45
			Sorg		
Av. yield loss (district) (t/ha)	0.06	0.07	0.04	0.06	0.06
Av. production loss (district) (1,000t)	5.30	6.34	3.47	5.08	4.91
Av. production cost (district) (mil usd)	1.00	1.20	0.66	0.96	0.93
Av. yearly total production loss (1,000 t)	262.99	311.16	134.30	445.46	430.35
Av. yearly total cost (mil usd)	49.65	58.75	25.36	84.10	81.25
			Millet		
Av. yield loss (district) (t/ha)	0.09	0.11	0.06	0.09	0.08
Av. production loss (district) (1,000t)	5.96	7.23	4.40	5.92	5.61
Av. production cost (district) (mil usd)	1.04	1.27	0.77	1.04	0.98
Av. yearly total production loss (1,000 t)	289.48	351.53	185.12	536.65	508.72
Av. yearly total cost (mil usd)	50.74	61.61	32.45	94.06	89.16
			Barley		
Av. yield loss (district) (t/ha)	0.01	0.02	0.05	0.04	0.03
Av. production loss (district) (1,000t)	0.12	0.28	0.57	0.42	0.32
Av. production cost (district) (mil usd)	0.02	0.05	0.11	0.08	0.06
Av. yearly total production loss (1,000 t)	2.57	6.09	12.00	18.08	13.78
Av. yearly total cost (mil usd)	0.49	1.16	2.29	3.45	2.63
			Maize		
Av. yield loss (district) (t/ha)	0.08	0.07	- 0.04	0.02	0.04
Av. production loss (district) (1,000t)	2.13	1.81	- 1.02	0.52	0.95
Av. production cost (district) (mil usd)	0.34	0.29	- 0.16	0.08	0.15
Av. yearly total production loss (1,000 t)	111.05	94.82	- 44.45	50.36	91.88
Av. yearly total cost (mil usd)	17.91	15.29	- 7.17	8.12	14.82

Notes:Individual cereal prices used represent the 2008 weighted cereal prices converted into USD using the average monthly exchange rate over this year. Aggregate cereal prices also use the 2008 cereal-specific prices. All numbers were rounded to two decimal places.

Table 5.9: Cost estimates - Dummies

	Main				
	Type 1 (only)	Type 1 (sep)	Type 2 (sep)	2 types (sep)	2 types (Together)
			Full sample		
Av. yield loss (district) (t/ha)	0.12	0.15	0.10	0.13	0.13
Av. production loss (district) (1,000t)	39.02	49.85	35.90	43.51	42.99
Av. production cost (district) (mil usd)	9.42	12.04	8.67	10.51	10.38
Av. yearly total production loss (1,000 t)	3,212.20	4,136.44	2,468.96	6,605.39	6,530.82
Av. yearly total cost (mil usd)	775.65	998.83	596.18	1,595.00	1,577.00
			Rice		
Av. yield loss (district) (t/ha)	0.16	0.20	0.15	0.17	0.17
Av. production loss (district) (1,000t)	23.79	30.38	22.52	26.79	26.34
Av. production cost (district) (mil usd)	7.12	9.10	6.74	8.02	7.89
Av. yearly total production loss (1,000 t)	1,691.33	2,165.23	1,351.88	3,517.11	3,480.88
Av. yearly total cost (mil usd)	506.51	648.43	404.85	1,053.28	1,042.43
			Wheat		
Av. yield loss (district) (t/ha)	0.10	0.13	0.09	0.11	0.11
Av. production loss (district) (1,000t)	10.87	14.08	10.88	12.59	12.51
Av. production cost (district) (mil usd)	2.43	3.15	2.43	2.81	2.80
Av. yearly total production loss (1,000 t)	637.44	824.96	557.78	1,382.74	1,377.14
Av. yearly total cost (mil usd)	142.54	184.47	124.72	309.19	307.94
			Sorg		
Av. yield loss (district) (t/ha)	0.05	0.06	0.02	0.04	0.04
Av. production loss (district) (1,000t)	5.06	5.83	1.85	4.07	3.96
Av. production cost (district) (mil usd)	0.95	1.10	0.35	0.77	0.75
Av. yearly total production loss (1,000 t)	247.75	286.11	71.40	357.51	347.50
Av. yearly total cost (mil usd)	46.78	54.02	13.48	67.50	65.61
			Millet		
Av. yield loss (district) (t/ha)	0.08	0.10	0.04	0.07	0.07
Av. production loss (district) (1,000t)	5.37	6.52	2.62	4.71	4.80
Av. production cost (district) (mil usd)	0.94	1.14	0.46	0.82	0.84
Av. yearly total production loss (1,000 t)	260.74	316.69	110.08	426.78	435.03
Av. yearly total cost (mil usd)	45.70	55.51	19.29	74.80	76.25
			Barley		
Av. yield loss (district) (t/ha)	0.01	0.03	0.05	0.04	0.04
Av. production loss (district) (1,000t)	0.12	0.30	0.56	0.43	0.42
Av. production cost (district) (mil usd)	0.02	0.06	0.11	0.08	0.08
Av. yearly total production loss (1,000 t)	2.46	6.39	11.96	18.35	18.06
Av. yearly total cost (mil usd)	0.47	1.22	2.28	3.50	3.45
			Maize		
Av. yield loss (district) (t/ha)	0.07	0.05	- 0.10	- 0.02	- 0.02
Av. production loss (district) (1,000t)	1.65	1.09	- 2.53	- 0.56	- 0.56
Av. production cost (district) (mil usd)	0.27	0.18	- 0.41	- 0.09	- 0.09
Av. yearly total production loss (1,000 t)	86.38	57.36	- 111.07	- 53.71	- 53.72
Av. yearly total cost (mil usd)	13.93	9.25	- 17.91	- 8.66	- 8.66

Notes:Individual cereal prices used represent the 2008 weighted cereal prices converted into USD using the average monthly exchange rate over this year. Aggregate cereal prices also use the 2008 cereal-specific prices. All numbers were rounded to two decimal places.

district at 160 kg/ha and 110 kg/ha for hot and cold droughts, respectively. These smaller impacts on yields translate into lower total economic costs. Whereas we estimate that, in a given year, the total economic cost of a hot drought is, on average, approximately USD 1.02 billion (using 2008 crop prices; column 2 in Table 5.8)[17], this falls to USD 650 million for a cold drought (column 3 in 5.8).

To our knowledge, there is only one study in the literature that attempts to estimate drought costs for the whole of the country. Using data for all Indian districts, Sarkar (2011) estimates cereal losses of 27.6 million tonnes due to drought in 2002[19]. In our sample of 275 districts (out of 311 apportioned districts), we estimate total production losses of about 16.3-16.9 million tonnes in 2002, which we value at 103 billion rupees using nominal prices. In addition to covering fewer districts, the differences in estimates are also likely to stem from methodological differences as well as the fact that we do not take into account a potential reduction in cultivated area during a drought year[20].

Omitting cold droughts can lead to a lower estimate of hot drought impact, especially when the dummy variables are excluded. Including dummy variables allows for a convergence in the marginal effect. However, if the change in intercept is taken into account when estimating costs, the divergences in the costs persist despite the inclusion of dummy variables. These effects are quantifiably large as we illustrate by comparing the first two columns of Tables 5.8 and 5.9 for the full sample. When dummies are excluded (Table 5.8), this effect is large. Average yield losses are estimated to be 33% higher (from 120kg/ha to 160 kg/ha) when cold droughts are included. These estimates have a substantial effect on the estimated average annual cost. This is USD 805 million (Table 5.8, column 1) when cold droughts are omitted compared to USD 1.02 billion (Table 5.8, column 2) when they are included, which represents a 27% increase. Thus, if estimating the economic cost of hot droughts without accounting for

[17]Crop prices in Indian rupees are converted into USD using the average monthly exchange rate obtained from x-rates.com[18]. More details on how prices are computed are available in Appendix C.

[19]The author values these losses at around 1.3 trillion rupees. An error in their calculations, however, suggests a loss closer to 130 billion rupees.

[20]A more detailed explanation is given in Appendix C. There are two further studies that attempt to estimate drought costs in India, which are less relevant for the purpose of comparison with our estimates. Pandey et al. (2007) estimate costs from yield losses for three states in Eastern India and the EM-DAT database bases its national-level cost estimates on the basis of losses in housing, agriculture and livestock. Within range of our estimates, the former estimate a cost of USD 900 million for the 2002 drought. However, a large number of droughts do not have associated costs. Specifically, for our sample period, only 12 drought events were recorded for India in the EM-DAT database and, out of these 12 events, only five have had their associated costs estimated.

cold droughts, the average yearly total costs of drought would approximate USD 805 million. Including cold droughts raises the cost by 107% to USD 1.67 billion (column 4 in Table 5.8). The difference can be broken down as follows: USD 216 million can be attributed to the lower coefficient of hot droughts, and USD 649 million to the inclusion of cold droughts. A similar difference exists when dummy variables are included, in Table 5.9[21].

The impacts derived using the separate crop-specific specifications generate patterns of yields and costs similar to those that emerge from our aggregate cereal specification. Also, the patterns that emerge in the full sample specification are present in the crop-specific regressions. Cold droughts have a particularly large impact on rice, leading to substantive physical and economic losses. Indeed, about half of the total average economic cost from cold droughts can be explained by losses in rice yields. In the absence of dummies, we estimate rice yield losses of around 160kg/ha (column 5 under Rice in Table 5.9). Excluding Type 2 drought, i.e. comparing column 1 and 2 under Rice in Table 5.9, suggests underestimates in the region of about 33%. Thus, it is highly likely that Birthal et al. (2015) underestimated the impacts of droughts on rice yields in their analysis[22].

5.7 Conclusion

Overall, there are three main findings that emerge from our analysis. First, after proposing an index which extends Yu and Babcock's (2010) index, we show that both hot and cold droughts have significant impacts on agricultural productivity in India. Thus, it is important to include the latter category of droughts, especially in a setting where there has been a clear increase in the number of such events in recent years. Moreover, if an assessment of economic impacts is performed solely based on hot droughts alone, approximately half of all potential dry events would be overlooked. Our results strongly suggest that these events have had quite a severe impact on cereal yields.

Second, the omission of cold droughts leads to a smaller estimated coefficient of hot droughts,

[21] However, in the case of Table 5.9, this difference arises from accounting for the intercept change, rather than the underestimation of the marginal impact.

[22] Birthal et al. (2015) estimate rice yield losses due to drought ranging from 187 to 200 kg/ha. Differences in estimated impact are likely to stem from the fact that they use a different sub-sample of districts and estimate a specification that differs from the one used in our analysis, e.g. we adopt a district-specific quadratic trend whereas they adopt a linear trend, as well as including interaction terms and irrigation as a control variable.

especially when a dummy variable is not included to account for a potential intercept shift. Effectively, this implies that, if cold droughts have a large negative effect on productivity, estimating the coefficient for hot droughts without accounting for cold droughts could lead to underestimates of the marginal effect of hot drought thus further downward biasing the overall impact of drought in empirical analyses. This result does not challenge the central findings of Yu and Babcock (2010) and Birthal et al. (2015) that impacts of drought have decreased over time. Yet, it does question the size of the marginal impacts estimated in both of these studies, and implies that a focus on hot droughts alone does not tell the whole drought story.

Third, the quantitative implications of our results are likely to be large, particularly given the fact that our cost estimates are based purely on yield losses. Since we do not take any potential changes in the cultivated area into account, we are likely to underestimate true production losses. The economic value of production loss attributable to cold droughts is illustrated to be approximately 60% of the total economic value of production losses attributable to hot droughts in our main specification. Also, omitting cold droughts and a dummy variable can lead to an underestimation of the economic value of production losses due to hot droughts, which in our simulations amounted to a difference of about 27%. While we acknowledge that our back-of-the envelope estimates are based on a number of assumptions regarding prices and so forth, they do suggest that we have found sufficient empirical evidence and an economic rationale to justify the inclusion of cold drought in analyses of drought impact.

Our results have clear implications for public policy. Since cold droughts have measurable impacts on agricultural production that are severe yet not as severe as those resulting from hot droughts, policymakers should seek to distinguish between the two types of drought defined in this paper. Simple metrics of precipitation deficiency will obviously capture both types but since temperature plays a critical role in determining the extent of dry conditions at the local scale, it still needs to be explicitly accounted for. Detecting cold drought and tracking their impacts over time can serve as an early-warning response for periods when temperatures are expected to be above the average of long-term trends. With global warming expected to continue to contribute to rising temperatures as well as potentially influencing patterns of extreme rainfall events, our index can thus help to shape the appropriate policy response to drought, particularly with respect to climate adaptation and agricultural production in more climate vulnerable locations.

References

[1] Auffhammer, M., Ramanathan, V. & Vincent, J. R. (2012). Climate Change, the Monsoon, and Rice Yield in India. *Climatic Change*, 111(2):411-424.

[2] Birthal, P., Negi, D., Khan, M., & Agarwal, S. (2015). Is Indian Agriculture Becoming Resilient to Droughts? Evidence from Rice Production Systems. *Food Policy*, 56, 1-12.

[3] Burgess, R., Deschenes, O., Donaldson, D. & Greenstone, M. (2014). The Unequal Effects of Weather and Climate Change: Evidence from Mortality in India.

[4] Diffenbaugh, N., Swain, D. & Touma, D. (2015). Anthropogenic Warming Has Increased Drought Risk in California. *Proceedings of the National Academy of Sciences of the United States of America*, 112(13), 3931-3936.

[5] Gadgil, S. & Gadgil, S. (2006). The Indian Monsoon, GDP and Agriculture. *Economic and Political Weekly*, 41(47), 86-91.

[6] Government of India, (2009). Manual for drought management. Technical report, Ministry of Agriculture, Department of Agriculture and Cooperation, Government of India, New Delhi.

[7] Guiteras, R. (2009). The impact of climate change on Indian agriculture. *Manuscript, Department of Economics, University of Maryland, College Park, Maryland.*

[8] Hatfield, J. L., Boote, K. J., Kimball, B. A., Ziska, L. H. & Izaurralde, R. C. (2011). Climate Impacts on Agriculture: Implications for Crop Production. *Agronomy Journal*, 103(2), 351-370.

[9] ICRISAT, (2012). District Level Database Documentation. Technical report, ICRISAT-ICAR-IRRI Collaborative Research Project.

[10] IPCC, (2012). Field, C. B. (Ed.). (2012). *Managing the risks of extreme events and disasters to advance climate change adaptation: special report of the intergovernmental panel on climate change*. Cambridge University Press.

[11] Indian Metereological Department, Government of India (undated). Frequently Asked Questions: FAQ. Available at: here

[12] Jain, S. K. & Kumar, V. (2012). Trend analysis of rainfall and temperature data for India. *Current Science (Bangalore)*, 102(1), 37-49.

[13] Jayachandran, S. (2006). Selling labor low: Wage responses to productivity shocks in developing countries. *Journal of political Economy*, 114(3), 538-575.

[14] Mishra, A. K. & Singh, V. P. (2010). A Review of Drought Concepts. *Journal of Hydrology*, 391(1), 202-216.

[15] Lobell, D. B., Sibley, A., & Ortiz-Monasterio, J. I. (2012). Extreme heat effects on wheat senescence in India. *Nature Climate Change*, 2(3), 186-189.

[16] Lobell, D. B., & Gourdji, S. M. (2012). The influence of climate change on global crop productivity. *Plant Physiology*, 160(4), 1686-1697.

[17] Pai, D. S., Sridhar, L., Rajeevan, M., Sreejith, O. P., Satbhai, N. S., & Mukhopadhyay, B. (2014). Development of a new high spatial resolution (0.25 0.25) long period (19012010) daily gridded rainfall data set over India and its comparison with existing data sets over the region. *Mausam*, 65(1), 1-18.

[18] Bhandari, H., Pandey, S., Sharan, R., Naik, D., Hirway, I., Taunk, S. K., & Sastri, A. S. R. A. S. (2007). Economic costs of drought and rice farmers drought-coping mechanisms in eastern India. *Economic Costs of Drought and Rice Farmers Coping Mechanisms: A Cross-Country Comparative Analysis, International Rice Research Institute, Los Baos, Philippines*, 43-111.

[19] Parcell, J. L., & Cain, J. (2013). Drought Tolerance of Soybean Crops in Missouri. *In Paper for Southern Agricultural Economics Association (SAEA) Annual Meeting, Orlando, Florida, February.*

[20] Prasad, P. V. V., Staggenborg, S. A., & Ristic, Z. (2008). Impacts of drought and/or

heat stress on physiological, developmental, growth, and yield processes of crop plants. Response of crops to limited water: *Understanding and modeling water stress effects on plant growth processes*, 301-355.

[21] Prasanna, V. (2014). Impact of monsoon rainfall on the total foodgrain yield over India. *Journal of Earth System Science*, 123(5), 1129-1145.

[22] J. Sarkar (2011). Drought, its impacts and management: Scenario in India. In R. Shaw & H. Nguyen (Eds.), *Droughts in Asian Monsoon Region* (pp.67-86). Bingley, UK: Emerald Group Publishing Limited.

[23] Sarsons, H. (2015). Rainfall and conflict: A cautionary tale. *Journal of Development Economics*, 115, 62-72.

[24] Schlenker, W., & Roberts, M. J. (2009). Nonlinear temperature effects indicate severe damages to US crop yields under climate change. *Proceedings of the National Academy of sciences*, 106(37), 15594-15598.

[25] Shah, M., & Steinberg, B. M. (2013). *Drought of opportunities: contemporaneous and long term impacts of rainfall shocks on human capital* (No. w19140). National Bureau of Economic Research.

[26] Sharma, K. D. (2011). Rain-fed agriculture could meet the challenges of food security in India. *Current Science*, 100(11), 1615-1616.

[27] Srivastava, A. K., Rajeevan, M., & Kshirsagar, S. R. (2009). Development of a high resolution daily gridded temperature data set (19692005) for the Indian region. *Atmospheric Science Letters*, 10(4), 249-254.

[28] Vicente-Serrano, S. M., Lopez-Moreno, J. I., Beguera, S., Lorenzo-Lacruz, J., Sanchez-Lorenzo, A., Garca-Ruiz, J. M., Azorin-Molina, C., Moran-Tejeda, E., Revuelto, J., Trigo, R. & Coelho, F. (2014). Evidence of increasing drought severity caused by temperature rise in southern Europe. *Environmental Research Letters*, 9(4), 044001.

[29] Wilhite, D. A. (2000). *Drought: A Global Assessment*. Routledge, London.

[30] Wilhite, D. A., & Glantz, M. H. (1985). Understanding: the drought phenomenon: the role of definitions. *Water international*, 10(3), 111-120.

[31] Yu, T., & Babcock, B. A. (2010). Are US corn and soybeans becoming more drought tolerant?. *American Journal of Agricultural Economics*, 92(5), 1310-1323.

Chapter 6

Threshold effects of extreme rainfall events and the evolution of drought impacts on Indian agriculture

Abstract

Climate change is predicted to be associated with a rise in the intensity and frequency of extreme rainfall events. Such patterns are already observed in India, which is highly dependent on rain-fed and irrigated agriculture. This paper examines the hypothesis that large deviations, either positive or negative, from long-term average rainfall have significant and large impacts on crop production in India and then investigates the evolution of impacts over time. A precipitation-temperature index and a threshold regression approach are applied to meteorological and agricultural data to test for the existence of thresholds in the relationship between rainfall and agricultural production, over the period 1966-2009. For India, significant marginal negative impacts of rainfall deficiency are found above the threshold previously used by the Government to denote a drought. Impacts become much larger at deviations of about 25%, and even more dramatic at deviations in excess of 50%. Thresholds in arid and semi-arid areas occur at larger negative deviations from normal rainfall. At low negative deviations from normal rainfall we find smaller impacts in more arid areas compared to humid areas. However, at large deviations, impacts are higher in more arid areas. Sorghum, millet and maize have lower thresholds than rice, wheat and barley. Over time, we find reductions in drought impacts until the late nineties, though the trend seems to have reversed since the beginning of the millennium. This pattern is consistent across agro-ecological zones and crops. One possible explanation is the decrease in lagged precipitation in drought affected districts.

Keywords: Agriculture, India, Rainfall, Thresholds, Cereals

JEL classification: Q10, Q54, Q56, O13

6.1 Introduction

A warming climate is predicted to be associated with a rise in the frequency and intensity of extreme weather and climate events, with profound implications for both human society and the natural environment (Easterling et al., 2000; IPCC, 2012). While the impacts of extreme events have the potential to be non-linear, the policy response to such events, especially drought, has often been determined by arbitrarily defined thresholds (Willhite and Glantz, 1985). This is especially true of India, where, until recently its government defined moderate and severe meteorological droughts as years where the rainfall deficiency vis-à-vis the annual monsoon rain was between 26%-50% and above 50%, respectively. Since these thresholds are not intrinsically tied to tangible outcomes, they fail to characterize the level of deviations from average rainfall levels beyond which additional deviations directly translate into an impact that could be considered significant.

In India, where the summer monsoon contributes up to 85% of the country's rainfall, recent research by Singh et al. (2014) indicates that levels of monsoon rainfall have decreased since 1951. They also find that periods of heavy rainfall have become more intense while 'dry spells' (drought) have become less intense yet more frequent[1]. This study, along with other studies of observed Indian rainfall extremes (e.g. Goswami et al., 2006; Pai et al., 2011; Ghosh et al., 2012; Turner and Annamalai, 2012; Kumar et al., 2013), all suggest critical implications for Indian agriculture. Given that the amount of rainfall is a key driver of agricultural productivity[2], we go a step further and examine the hypothesis that large deviations, either positive or negative, from long-term average rainfall have significant and large impacts on crop production. We do so by using meteorological and agricultural data, collected in India between 1966 and 2009, to test for the existence of thresholds in the relationship between rainfall and agricultural production, before estimating the magnitude of impact on productivity and frequency of extreme rainfall events over time.

Extreme rainfall has the potential to lead to substantial welfare costs for producers and consumers, through, respectively, lost income and higher food prices. India, with an agricultural

[1]They define dry and wet spells as "events of at least 3 consecutive days with precipitation anomalies consistently exceeding one standard deviation of daily precipitation".

[2]The production of crops in many areas during the wetter summer (Kharif) season relies directly on rainfall as their main source of water. Crops grown in the subsequent drier (Rabi) season also rely on rainfall from the previous season for soil moisture and water stored in sources such as tanks and canals.

sector contributing about 20% of gross domestic product (GDP) and employing half of the working population, is particularly vulnerable to large changes in precipitation. Two-thirds of the country is susceptible to drought (Birthal et al., 2015), and rain-fed agriculture covers approximately 60% of its cropped area (Sharma, 2011). About 40 million hectares are considered flood-prone (De, Dube and Rao, 2005). Previous research on impacts of extreme rainfall events on Indias economy and agricultural sector has mostly focused on dry spells. Between 1951 and 2003, severe droughts were estimated to have lowered the countrys GDP around 2 to 5 percent (Gadgil and Gadgil, 2006). Closest to our study is Singh et al. (2011) who find that, in some cases, excess rainfall had a negative impact on aggregate agricultural production, although the main impacts were from deficient, rather than excessive, rainfall. However, they define extreme rainfall events and estimate their impacts arbitrarily and they do not take into account temperature when analysing deficient rainfall events.

In this paper, we seek to address two questions. First, using a simple and flexible precipitation-temperature index, which interacts rainfall with temperature, we estimate the thresholds of rainfall and their impacts on cereal production for different ranges of rainfall. Second, focusing solely on deficient rainfall, we investigate how impacts of drought have changed over time. For the first question, a threshold regression approach (Hansen 1999, 2000) is applied to our panel of district-level agricultural data. Together, the index and the threshold approach allows us to identify data-driven ranges along which impacts of extreme rainfall events have significant impacts on agricultural production, while simultaneously accounting for the potential effect of temperature on these impacts.

Conceptually, we argue that our approach addresses the main weakness of definitions of extreme rainfall events adopted in previous work, namely that they are based on arbitrary thresholds of excess rainfall or rainfall deficiency. The main benefit is that it ties deviations in rainfall from the long-term average to its effects on agricultural productivity. It is also sufficiently flexible in that it allows us to incorporate the effects of temperature, which have been shown to reduce yields of major crops in various settings (Schlenker and Roberts, 2009; Lobell et al., 2012; Deryng et al., 2014). Additionally, the method has the potential to capture impacts of both excessive and deficient precipitation on agricultural yield.

With regards to the second question, we first use a rolling-regression approach to analyse the evolution of drought impacts without any prior assumptions. Then, we use a standard fixed-effects panel data model to estimate how impacts of drought evolve over time. Our approach with regards to the evolution of impacts over time also addresses some of the weaknesses of previous papers in this literature which typically impose linearity of impacts over time and have generally concluded that impacts have become smaller over time (Birthal et al. 2015).

The paper proceeds in four parts. First, we identify thresholds for India as a whole. Second, we examine the data for any evidence of homogenous thresholds across different agro-ecological zones and crops. Disaggregating thresholds at a reduced geographical scale allows us to consider the heterogeneity in bio-physical conditions across India. For example, the strong possibility that arid areas may react to large changes in rainfall very differently in comparison to humid areas. Similarly, different crops may have very different sensitivities to deviations in rainfall. Previous studies focusing on temperature (e.g. Schlenker and Roberts, 2009) have shown different thresholds depending on the crop analysed and a similar pattern is expected to emerge for rainfall in our analysis, the first to our knowledge, to identify thresholds for precipitation. Third, we turn to the question of how the impacts have evolved over time and we estimate a number of rolling regressions at a reduced geographical scale and by crop. This allows us to assess whether the evolution of the impacts of drought over time has been different across agro-ecological zones and crops. Finally, after testing for the most appropriate parametric specification, we estimate a parametric fit of the impacts of drought over time.

A number of findings emerge from our analysis. The first is that the vast majority of thresholds are thresholds of rainfall deficiency. For India as a whole, we find significant small marginal negative impacts of rainfall deficiency up until a negative deviation of approximately 25% from long-term average rainfall (LTAR). Beyond a 25% deviation from long-term average impacts become larger and become even more dramatic at rainfall deviations in excess of 50%. Across agro-ecological zones there is spatial heterogeneity in impacts. Thresholds in arid and semi-arid areas tend to occur at larger negative deviations and impacts at small deviations tend to be smaller than in sub-humid areas. However, at large negative deviations, we find very large impacts for arid areas. Our analysis by crop finds that sorghum, millet and maize (to a lesser extent) tend to have lower thresholds, whereas rice and wheat have higher thresholds. This implies that crop choice is likely to be an important component of resilience to climate change. However, the fact that the majority of thresholds are thresholds for deficiency does not mean that excessive rainfall cannot have negative impacts. However, these were confined to restricted agro-ecological zones and crops and, in most cases, negative impacts only occurred at extreme amounts of rainfall and impacts were quantitatively small.

Finally, in terms of the results over time, we find that, overall, impacts of drought over time in India have been highly non-linear and, while impacts did become smaller until the late nineties, they seem to have become more severe since the beginning of the millennium. However, the results by crop are highly heterogeneous, with rice witnessing a high and steady decrease in impacts over time, whereas impacts on maize and sorghum have been more non-linear over time. We argue that the differential pattern by crop is likely to be driven by factors such as the adoption of irrigation and improved varieties. One possible explanation to the worsening of impacts since the beginning of the millennium one possible explanation is what seems to be a change in rainfall patterns. In our data, we observe that the rainfall received in the year prior to the drought was particularly low since the early 2000s. As a result, the lack of availability of moisture from the previous year could explain why we observe this increase in impacts.

Where reliance on agricultural income is high, extreme rainfall events may have devastating effects on human welfare and pose a significant challenge to policy-makers who manage the response to these events. Future projections of climate change-induced changes in rainfall patterns across India suggest increasingly erratic rainfall. Our results therefore contribute to a better understanding of the vulnerability of agriculture to extreme rainfall events given the critical need for food security in the face of the growing threat of climate change. The rest of the paper is structured as follows. Section 2 provides background on India and discusses measurement issues. Section 3 discusses the data and methodology used in this paper. Section 4 presents and discusses the results. Finally, section 5 concludes.

6.2 Agriculture in India and measuring extreme rainfall events

Deviations in rainfall from its long-term average play an integral role in Indian agriculture and its occurrence has large implications for the economy as a whole. The impacts of extreme rainfall events in India are conditioned by a number of factors that vary across the country.

Given the large size of India, it is debatable whether we are able to properly characterise drought impacts based on a country-wide average. The rest of this section first reviews sources of heterogeneity that may affect the impacts of rainfall deviations before examining how extreme rainfall events have been measured in previous research.

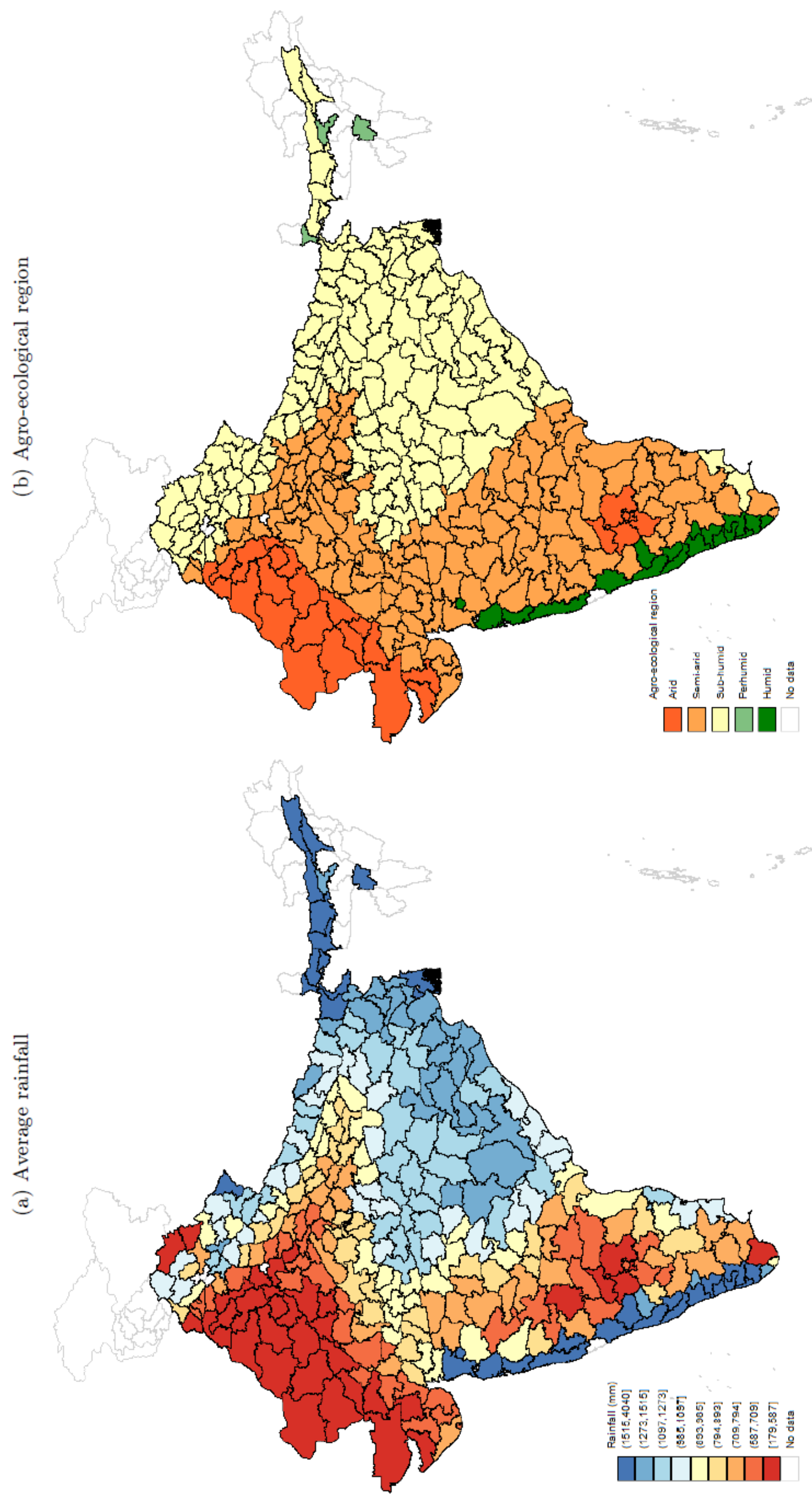
6.2.1 Agro-climatic differences

Climatic conditions vary substantially across growing regions. This is illustrated in Figure 6.1. Panel (a) shows average levels of annual rainfall in each district across the country. Areas in the north-west of the country are characterised by extremely low average rainfall, in contrast to areas in the east and coastal-west that have much higher levels of rainfall on average. These differences in mean rainfall are primary determinants of a permanent feature of regions: aridity.

Estimating extreme rainfall impacts separately for these different zones is important for a number of reasons. First, identifying areas of drought vulnerability based on climatic differences is important for informing policy about future vulnerability. If regions already frequently exposed to dry conditions are most affected by drought, it is likely that future warming could exacerbate those already challenging growing conditions. Second, understanding the difference in sensitivity to rainfall deviations can help policy-makers identify when a drought is likely to start harming agricultural productivity. Given that, for instance, arid areas experience generally low levels of absolute rainfall, it may be simplistic to assume that a given proportional deviation below average would have effects on agricultural productivity comparable to an area with very high absolute levels of rainfall. For instance, a 20% deviation in rainfall from the long term average would amount to 30mm in arid areas, while the same proportional deviation would be around 200mm in humid areas. This may have substantially different effects on crop growth in these areas.

Physical exposure to drought may vary substantially across the country. Thus, we divide India into distinct regions based on their average agro-climatic characteristics. Panel (b) in Figure 6.1 shows a characterisation of Indian districts based on similar agro-climatic factors. Prior research suggests that India can be split into twenty agro-climatic regions based on a number of climatic variables, such as rainfall and temperature, and soil characteristics (Gajbhiye and

Figure 6.1: India rainfall by AEZ



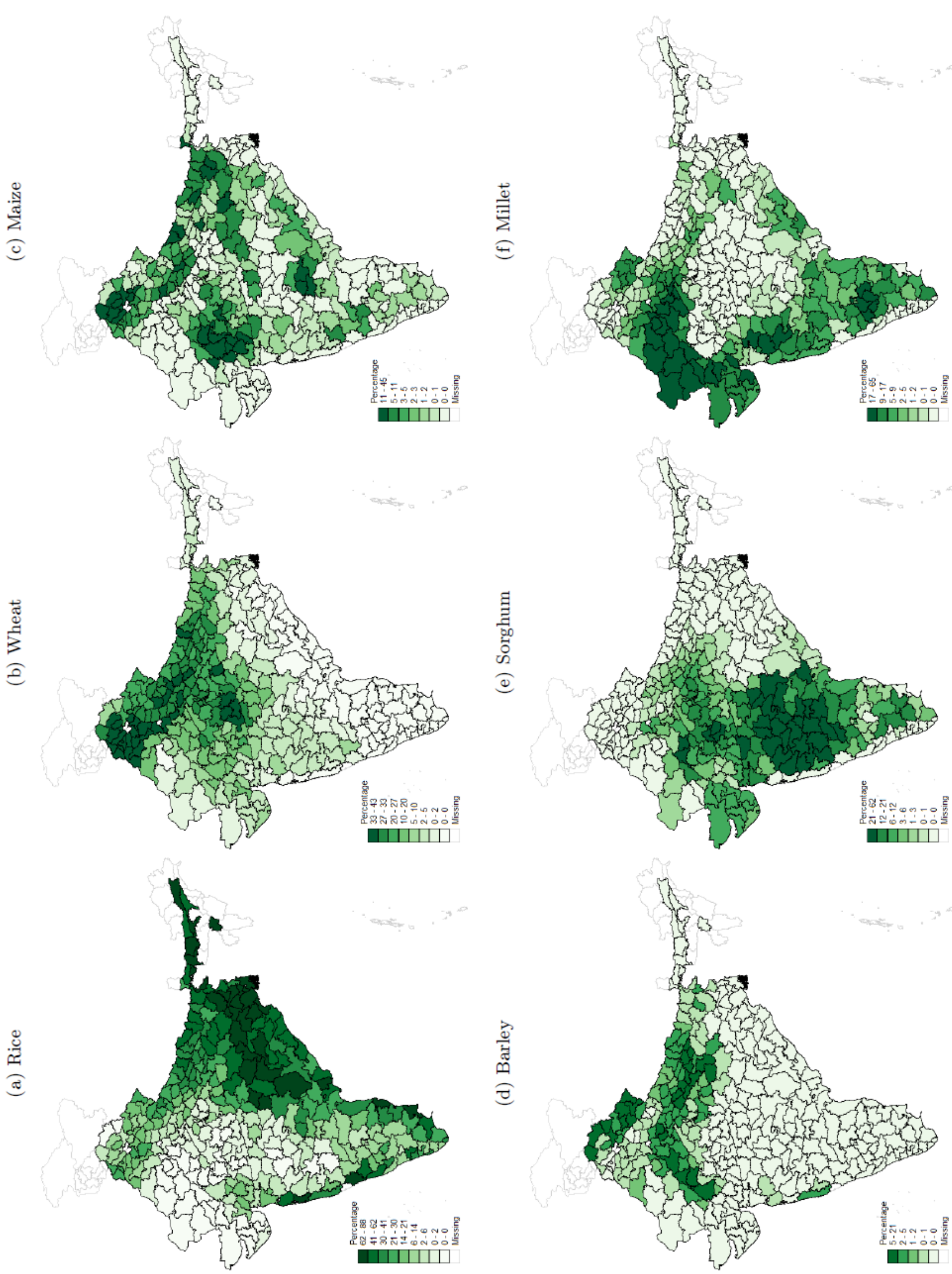
Note: Panel (a) shows district-wise average rainfall over the period 1957-2009. Panel (b) maps the agro-ecological grouping of districts using the methodology described in this section. For both maps, only districts with available agricultural data are shaded. Districts with no data are shown as white polygons. District boundaries refer to those drawn in 1966.

Mandal, 2010). We use this classification of agro-ecological zones to group districts into a broader classification depending on whether districts fall into arid, semi-arid, sub-humid, or humid zones. This allows us to maintain a relatively large number of districts in most zones to aid the empirical analysis. It can be seen by comparing panels (a) and (b) that this classification of zones corresponds very clearly with patterns of average rainfall, indicating that average rainfall is one of the main driving factors behind the variation in agro-climatic conditions across the country.

6.2.2 Crop choice

Another aspect that may determine the impacts of extreme rainfall events is crop type. Given the variation in average climatic conditions shown in Figure 6.1, crop choice in a district is likely to reflect these conditions. For instance, water-intensive crops are more likely to be grown in less arid areas. Figure 6.2 shows the spatial distribution of the proportion of area planted with each of the six crops examined in this study. Rice is planted most intensively in areas with high rainfall in the south and east of the country, where conditions are semi-humid or humid. In contrast, wheat is grown mainly in the more arid northern part of the country, which reflects a lower dependence on rainfall. The crops most suited to growth in dry environments, sorghum and millet, are both grown across arid and semi-arid regions. Similar to research that has shown how different crops react differently to heat stress (Schlenker and Roberts, 2009), different crops may also differ in terms of their resilience to water stress (FAO, 2012). The differences in levels of rainfall deviations for which changes in rainfall have a negative significant impact on productivity is considered in our empirical analysis.

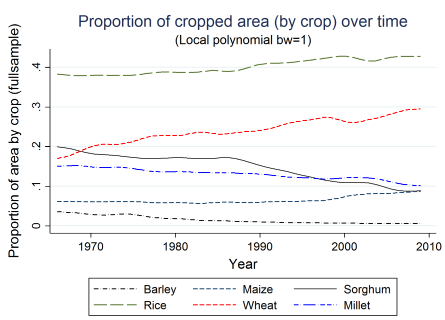
Figure 6.2: Area Planted by crop



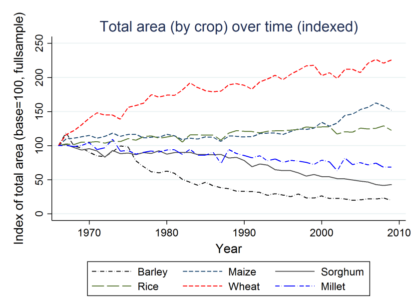
Note: Panels (a)-(f) show the proportion of gross cropped area devoted to each of the six crops used in the analysis. For districts with no data, these areas are shown as white polygons. District boundaries refer to those in 1966.

In addition to the natural differences between crops, we also note that, over time, the growth of cultivated area of certain crops as well as technology adoption (e.g. irrigation, fertilizer and high-yielding varieties) has been very heterogeneous. This is highlighted in Figure 6.3 which show that for districts in our sample there has been a shift away from sorghum, barley and millet towards rice, wheat and, to a certain extent, maize [figure 6.3 panels (a) and (b)]. With respect to technology adoption, crops such as wheat have benefited from the highest increase in irrigation [figure 6.3 panel (d)] and, while adoption of improved varieties increased across all crops, rice and wheat were the two crops with the highest area under high-yielding variety cultivation [figure 6.3 panel (c)]. These factors are likely to have an effect both on the location (and impacts) at a given threshold, but also affect how the impacts evolve over time.

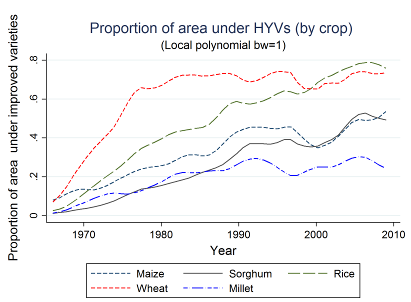
Figure 6.3: Area cultivated, adoption of high-yielding varieties and irrigation (by crop)



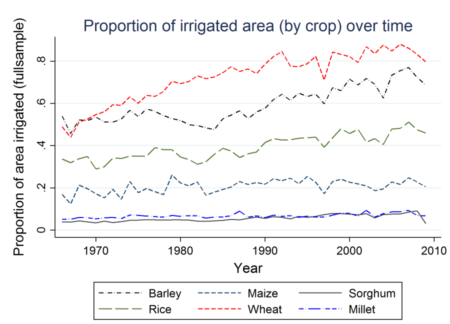
(a) Crop area (% of total)



(b) Total crop area (indexed)



(c) HYV crop area (% crop area)



(d) Irrigated crop area (% crop area)

Notes: panels (a), (c) and (d) are local polynomials (bandwidth= 1). Panel (b) refers to the relative total cultivated area for a given cereal, where, for the reference year (1966), all the series take a value of 100. For the HYV (high-yielding varieties) data, there were a number of occurrences where, for some districts, area under HYV cultivation exceeded the total area under cultivation. In these cases, these values were simply replaced by a missing value. This was an issue for about 10% of the observations. Given how noisy the HYV data are, we will use them for the rest of the analysis. However, they illustrate the rapid adoption of HYVs in India during the 1966-2009 period.

6.2.3 Measuring the severity of extreme rainfall events

There is no universal definition of the conditions that constitute a flood or a drought. In the case of drought, it is generally referred to as an extreme natural event associated with water deficiency over an extended period of time (Mishra and Singh, 2010). The severity of a drought and its impacts, however, are determined by a number of other factors, natural and man-made, which may differ substantially across space and time. Thus, a wide range of indices have been developed and used in research and policy. These range from simple precipitation indices, which are highly favoured among policy-makers, to very data-intensive multidimensional measures. In the case of floods, most of the indices used in the empirical literature tend to rely solely on rainfall indices vis-à-vis long-term average rainfall (Singh et al., 2011, Aufhammer et al., 2012).

A number of studies have estimated impacts of rainfall deficiency on agricultural production using simple metrics of precipitation deficiency. These measures have the advantage of being easily interpretable and capture the most obvious characteristic of drought, rainfall deficiency. For instance, a commonly used method is similar to that of Pandey et al. (2007) who define drought and severe drought as rainfall between 70-80% and below 70% of LTAR, respectively. They use this definition to estimate drought impact in areas that grow rice in Asia, at the aggregate and household level. In Eastern India, the authors find that drought is associated with a 36% loss of production value in Eastern India. A similar definition is also used by Auffhammer et al. (2012) to study the effect of monsoon rainfall on rice yields for states in India. They define drought if monsoon rainfall is 15% below normal and find that drought was associated with a 12% fall in rice yield. Singh et al. (2011) define drought as an event where the rainfall deficiency was more than one standard deviation below the mean and find significant negative impacts on cereal production.

In the case of excess rainfall, empirical research tends to be more limited. Aufhammer et al. (2012) define extreme rainfall as the summed June-September rainfall that occurred on days with rainfall that equalled or exceeded a states 95th percentile daily threshold. The authors find that while yield decreased by 0.2% for a 1% decrease in the cumulative June-September rainfall, the negative effect of an increase in 1% of extreme rainfall was only 0.022%. Similarly, Singh et al. (2011), who define wet years as those where monsoon rainfall exceed the long-term average by more than a standard deviation, only find negative impacts in two out of nine wet years.

Studies that use simple definitions of extreme rainfall events are problematic for our understanding of their impacts since they impose arbitrary thresholds in order to define a drought or flood, and evaluate their impacts only after a given level of precipitation. For drought, it is not clear whether such thresholds have an agronomic or empirical basis (Wilhite and Glantz, 1985). Also, there may be other variables that may have important effects when determining the physical severity of a drought, in particular temperature. High temperatures have acute effects on crop growth during periods of low precipitation since the rate of evapotranspiration, the combined process of water evaporated from land surfaces and plants, increases as temperatures rise (Prasad et al., 2008; Lobell and Gourdji, 2012). In general, this increases a plant's demand for water at a time when water availability is already reduced due to deficient precipitation.

The role of temperature in determining the physical severity of drought is also critical given temperature increases driven by climate change (Hatfield et al., 2011). Recent research has documented that droughts over a range of settings have increased in severity as mean temperatures have risen. Higher temperatures, rather than the increased intensity of low rainfall events, have been responsible for these drying trends (Vicente-Serrano et al., 2014; Diffenbaugh et al., 2015). Empirically, high temperatures have been shown to have detrimental effects on crop yields. Schlenker and Roberts (2009) find that temperatures reduce county-level yields for corn, soy-beans, and cotton in the U.S. In India, Guiteras (2009) and Burgess et al. (2014) both show that, on average, daily temperatures above 34°C tend to reduce agricultural productivity of a district. Lobell et al. (2012) identify the same threshold as harmful for wheat yields in India. A failure to consider the effect of temperature on the severity of a drought event could lead to a serious underestimation of its severity and give misleading information about the likelihood of future production losses driven by climate change.

6.2.4 Measuring impacts of drought over time in India

In an attempt to assess whether droughts have become more severe over time, Birthal et al. (2015) apply the methodology proposed by Babcock and Yu (2010) to the same dataset we use. They focus exclusively on rice and find, in India, a pattern similar to the one found in Babcock and Yu (2010), namely a reduction in impacts over time. The authors attribute a large part of this reduction in impacts to improvements in irrigation that have occurred during the Green Revolution. The finding that the impacts of drought have declined over time is of high importance in the context of India. However, we argue there are three issues with the authors analysis.

The first is a conceptual issue which comes from their use of a slightly modified version of the Babcock and Yu (2010) index, which we will discuss in the next section. The implication is that the authors definition of drought is limited to events characterized by both below-average rainfall and above-average temperature. We argue that this is an incomplete characterization of drought, given that below-average rainfall can still have largely negative impacts, despite below-average temperature.

The second issue relates to the focus of the authors on rice. Their focus is understandable since rice remains one of the most important crops in India. However, the narrow focus on rice may be problematic for two reasons. First, focusing on rice is particularly relevant for sub-humid and humid areas (and some semi-arid areas). As shown clearly by figure 6.2, rice is primarily cultivated in Southern and Eastern India. However, in Western and Northern India, wheat, maize and sorghum are important crops. Second, as shown in the previous section, rice and wheat are the two crops that have benefited the most from adoptions of high-yielding varieties and irrigation. As a result, it is possible that the results that the authors found for rice do not extrapolate to other crops, some of which still represent a large proportion of land under cultivation.

Finally, the third issue is a methodological one. The authors impose linearity in the evolution of impacts over time, which may condition the results. There are reasons to believe that impacts of drought in India may be highly non-linear over time. Shah and Kishore (2009), for instance, mention that the 2002-2003 drought was particularly bad since it was preceded by two years of deficient precipitation, which implied that the surface reservoirs were nearly empty and the depleted groundwater reservoirs had no chance to recover. Whether imposing a linearity over time is defendable or not really depends on the question at hand. If we are concerned about very long-term trends, arguably it is defendable, although it should be tested.

However, in our case we believe that it is important to attempt to capture some of these recent changes in the impacts of drought. High impacts in the early of 2000s would highlight that despite large increases in irrigation and potential reduction of drought impacts over time, the Indian agricultural sector is still not drought-proof and immune to large negative shocks.

As such, we will approach the evaluation of impacts over time slightly differently. First, we will propose a new index which has the same benefits of the Babcok and Yu (2010) index while addressing what we perceive to be its main weakness, which is the topic of the next section. Second, we will not impose linearity *a priori*, opting for a more flexible rolling regression approach to inform our choice of parametric trend, an approach we discuss in the methodology section.

6.3 Data and Methodology

6.3.1 Data

Our agricultural data are taken from the ICRISAT Meso-level Database, which contains information on a range of agricultural and socioeconomic variables at the district-level (ICRISAT,2012)[3]. We use data for the years 1966-2009. Out of the 311 available districts in the database, 275 districts are used to create a balanced panel for the main analysis[4].

Data are available on annual crop production and area, which are used to construct crop yield variables for rice, wheat, maize, barley, sorghum, and millet[5]. We investigate extreme rainfall impacts on an aggregate cereal productivity index as well as separately for each crop. The aggregate cereal productivity index is constructed by summing total cereal production and dividing this by total cereal area. Data on area irrigated and fertiliser consumed in a district are also used. These variables are available as district-level aggregates and are not crop-specific. In addition, socioeconomic census data are available at the district level.

[3]Since 1966 a number of districts have split into smaller districts. To maintain spatial consistency over time, district splits are dealt with by returning split districts to their parent districts in 1966.

[4]We test a number of different specifications. For India as a whole the number of districts ranges from 212 to 275, depending on the specification. All the specifications used require a balanced panel.

[5]For millet we add data on quantities of pearl millet and finger millet to create an aggregate quantity of millet.

To construct the precipitation-temperature index (discussed further in the next subsection), we use weather data for daily rainfall and daily average temperatures collected by the Indian Meteorological Department. The rainfall data are available in gridded format at a resolution of $0.25° \times 0.25°$ (Pai et al., 2014). Gridded temperature data are available at a resolution of $1° \times 1°$ (Srivastava et al., 2009). District-level weather data are then obtained by taking a weighted average of gridded weather observations from grid cells that fall within a district's boundary based on the proportion of the grid cell that falls in each district. Table 6.1 provides summary statistics of some of key variables.

Table 6.1: Summary statistics

Variable	N	Mean	S. D.	Min.	Max.
Yield					
Cereal yield (t/ha)	12100	1.463	0.787	0.006	4.775
Barley yield (t/ha)	5842	1.383	0.688	0.048	5.400
Maize yield (t/ha)	10621	1.431	0.929	0.003	9.739
Millet yield (t/ha)	9852	0.798	0.439	0.000	4.000
Rice yield (t/ha)	11398	1.492	0.853	0.009	5.542
Sorghum yield (t/ha)	9694	0.774	0.434	0.001	9.836
Wheat yield (t/ha)	10275	1.643	0.878	0.046	6.324
Area					
Cereal Area (1,000,000 ha)	12100	0.332	0.195	0.001	1.334
Barley (% of total district cereal area)	12031	0.015	0.036	0.000	0.320
Maize (% of total district cereal area)	12099	0.065	0.113	0.000	0.838
Millet (% of total district cereal area)	12100	0.131	0.220	0.000	1.000
Rice (% of total district cereal area)	12100	0.401	0.357	0.000	1.000
Sorghum (% of total district cereal area)	12066	0.148	0.225	0.000	0.929
Wheat (% of total district cereal area)	12093	0.240	0.246	0.000	0.972
Inputs and weather					
Irrigation (% of total district area)	12095	0.355	0.270	0.000	1.467
Rural population density	11787	3.566	2.142	0.428	17.907
Fertiliser intensity (t/ha)	11889	60.571	61.406	0.000	614.493
Cumulative rainfall (mm) (June-September)	12100	863.837	529.348	13.125	5313.428
Degree Days (cumulative heat, June-September)	12100	94.422	47.204	2.697	278.413
NRTI (rainfall-temperature index)	12100	0.260	0.202	0.000	1.000

Notes: N refers to the total number of observations. S. D. refers to the standard deviation. Min. and Max. refer to the minimum and maximum values. Rural population density is calculated as the total rural population divided by the gross cropped area. Fertilizer intensity is obtained by dividing total fertilizer used by gross cropped area. The cumulative rainfall variable is obtained by summing the cumulative precipitation in the months of June through to September. The degree-days variable refers to the total number of degree-days above the long-term average temperature during the growing season (defined as June-September). The NRTI variable is derived by multiplying the normalized negative rainfall by the normalized degree-days variable. More details are available in the data preparation appendix in the paper.

6.3.2 Rainfall-temperature index (RTI)

In this section, we develop a precipitation-temperature index. This expands the approach of Yu and Babcock (2010), who limit their definition of drought to years when both temperatures are uncommonly high and precipitation low relative to the long term average of these variables[6]. Since we remain agnostic about the response function of yield to our index, we simply interact our measure of precipitation with temperature. Specifically, we normalize the negative of our precipitation variable, such that the highest precipitation deficiencies take higher values, with a range of 0 to 1. Similarly, we normalize the degree-days of heat exposure between 0 and 1 such that hotter years have higher values. Then, we simply interact the two normalized values so that our precipitation-temperature interaction index ranges between 0 and 1, with values close to 1 signifying hot and dry years and values close to 0 signifying cold and wet years.

For values of rainfall above average rainfall for a given year, we expect a positive coefficient of our index on productivity (as rainfall approaches long-term average rainfall). However, for values of rainfall below a certain point of rainfall deficiency we expect a negative relationship (as rainfall deficiency increases)[7].

6.3.3 Methodology

Threshold regression with fixed effects

To estimate the impact of extreme rainfall events on Indian agriculture, we employ a threshold regression estimation strategy with fixed effects (Hansen, 1999)[8]. This model augments the standard linear fixed effects model by estimating how the effect of extreme rainfall on crop yield differs between thresholds of precipitation relative to its long-term average.

[6] We limit our analysis to considering drought as a prolonged absence of rainfall over the period from June-September. As such, we do not analyse, for instance, shorter or longer periods of drought. For instance, Fishman (2016) studies the intra-annual distribution of rainfall in India and concludes that this has important effects on productivity. To analyse the impacts of rare, multi-year droughts we would require a drought measure with 'memory' that takes into account soil moisture conditions. Since drought in India is mainly driven by variation in the annual monsoon, we argue that this measure is most relevant in this context.

[7] Conceptually, it is possible for us to have negative impacts of the index even for regions of rainfall above 1 since it could also signify high temperatures and/or that the average amount of rainfall may not be in the optimal region for the crop in question.

[8] To estimate the fixed effects threshold model we utilise Stata code which is described in Wang (2015)

Equation 6.1 illustrates the model in the case of a single threshold of precipitation (q_{it}), RT_{it}, the precipitation-temperature index (hereafter index) variable and $ln(y_{it})$, is the de-trended (quadratic district-specific trend) natural logarithm of crop yield[9]. De-trended crop yield is used since the threshold regression approach precludes the use of trended data and integrated processes. De-trending in this way removes trends in yields that are associated with technological progress over time[10]. To check if the yield variable is stationary after the de-trending procedure we apply a number of panel unit root tests[11]. In a number of specifications we also include a set of control variables X_{it}. Finally the error term is given by e_{it}.

$$ln(y_{it}) = \alpha_i + RT_{it}(q_{it} < \gamma)\beta_1 + RT_{it}(q_{it} > \gamma)\beta_2 + X_{it}\delta + e_{it} \tag{6.1}$$

which can be written as:

$$ln(y_{it}) = \alpha_i + RT_{it}(q_{it}, \gamma)\beta + X_{it}\delta + e_{it} \tag{6.2}$$

where:

$$ln(y_{it}) = \begin{cases} \alpha_i + RT_{it}\beta_1 + X_{it}\delta + e_{it} \text{ if } q_{it} < \gamma \\ \alpha_i + RT_{it}\beta_2 + X_{it}\delta + e_{it} \text{ if } q_{it} > \gamma \end{cases} \tag{6.3}$$

Rather than the effect of changes in our index being constant across all values of the threshold variable (ranges of rainfall, q_{it}), the threshold model estimates the value of one or more thresholds $q_{it} = \gamma$, for which the marginal effect of changes in our index has a different effect on cereal productivity. In other words, the estimated marginal effect of our index is different on either side of the estimated threshold. This method allows us to test whether such a

[9]The log transformation of yield is used because we are interested in the relative impact of an extreme event. This specification allows for a better comparison of impacts of extreme weather events across areas where absolute differences may be large.

[10]De-trending ensures that our results are not driven by these endogenous trends. If we do not de-trend, de-mean or take the first difference, the threshold may be driven by the unit root process followed by our series.

[11]We show the results of unit root tests in Tables 6A.1 and 6A.2. We also check the sensitivity of our results to alternative, stationary dependent variables, namely the de-meaned version of the natural logarithm of yields.

threshold exists and, if so, enables us to estimate threshold values and allows us to compute impacts for different ranges of precipitation.

The threshold value is estimated by least squares and involves picking the value that minimises the residual sum of squares of the model (Hansen, 2000). As argued by Hansen (1999), prior to searching for a threshold we need to eliminate the largest and smallest n% of the threshold variable (trimming). The remaining values of the threshold variable constitute the searchable values of a threshold. Another important feature of the model is that, even if a threshold is estimated, it may not be statistically significant. Accordingly, a likelihood ratio test of whether $H_0 : \beta_1 = \beta_2$ is implemented. A bootstrap procedure ran over 300 iterations is used to construct the p-values for this test. If we fail to reject H_0, the model is equivalent to the linear fixed effects model, where the effect of the regressors included in the model are not significantly different across values of rainfall. The method also allows us to compute a maximum of three threshold values.

A benefit of using panel data to measure extreme rainfall impacts is that it allows us to control for the influence of time-invariant factors that may differ among districts. The district fixed effect term, α_i, is included to control for time-invariant, district-specific effects, such as soil types or differences in altitude. Similarly, this could capture institutional differences that have persisted over the sample period. Such differences could help explain variability in extreme rainfall impacts on productivity across the country. In sum, identification of the impact of extreme rainfall events relies on within-district variation in de-trended yields, which exploits variability in the severity of impacts over time.

To test the robustness of our results, we also estimate our results with and without a set of time-varying control variables X_{it}. We include rural population per hectare of cereal area, total cereal area, fertilizer used and proportion of land under irrigation. Finally, we check whether the results of our threshold model are in line of the results we obtain when we estimate deviations of rainfall from the long-term average using dummy variables.

6.3.4 Rolling regressions and fixed effects model

To address the question of how impacts have evolved over time, we prefer to remain agnostic towards the potential shape of this relationship over time since, for the reasons highlighted previously, it has the potential of being highly non-linear. In order to explore the evolution of the coefficient over time, we first estimate our fixed effects model with fixed rolling windows. The main aim behind this method is to inspect how a given parameter varies over time by defining a window of a given width and then estimate the model rolling forward one period at a time. In our case, where the data runs from 1966 to 2009, we choose a 9-year window[12].

The estimates obtained from the rolling regressions is that they are likely to display a trend as well as some noise. As such, if the coefficients are constant over time, we should expect the coefficients to display random noise. However, if the coefficients are trending, we would expect this method to pick the true coefficient plus some noise. In our case, we use these estimates in order to potentially motivate the use of higher order time polynomials, rather than just interacting the drought coefficient with time linearly. In order to see if a consistent pattern emerges, we use different window widths[13].

Following the estimation of the rolling regression, we estimate models where drought is interacted with time trend. However, we allow the drought coefficient to be interacted with higher order polynomials of time. We estimate models up to an interaction of the drought index with a cubic interaction of time. In terms of a regression, we estimate, at most, the following regression[14]:

$$ln(y_{it}) = \alpha_i + \delta_{i1} * t + \delta_{i2} * t^2 + \theta D_{it} + \beta_0 DI_{it} + \beta_1 DI_{it} * t + \beta_2 DI_{it} * t^2 + \beta_3 DI_{it} * t^3 + e_{it} \quad (6.4)$$

[12]This means that first, we would estimate the model for 1966-1975 (included), then for 1967-1976, and so forth until 2000-2009. Then we can attempt to better understand the potential shape of the coefficient over time. In our case, for each window T from t to t+9, included, we estimate a fixed-effects model and we retrieve the drought coefficient.

[13]We run the rolling regressions using window widths of 3, 6 and 9 years. For the 3-year case, no trends were added to the rolling regression specifications. For all the remaining windows, we added quadratic, district-specific trends to the rolling regressions.

[14]Note: In the regression tables 6A.15-6A.20 the index is called $NRTIq_{12}$ because it is equal to the NRTI when the values of rainfall are below the LTAR.

We then use an F-test to determine which is the most appropriate parametric specification (i.e. whether there is support for a trend and, if so, of what order).

6.4 Results and Discussion

6.4.1 Threshold results, discussion and robustness checks

Full Sample

We begin by investigating precipitation thresholds for India as a whole (see Table 6.2 for full results; a summary of all our results and marginal effects can be seen in Table 6.3). Units of precipitation refer to the proportion of annual rainfall relative to a districts long-term average rainfall (LTAR). Our results suggest a small negative coefficient of our index for levels of rainfall above 74% of LTAR. Below 74% of LTAR, however, we find significant and sizeable negative impacts, with an increase in the index leading to an additional estimated -0.266% deviation of yields from trend; these negative impacts become even more pronounced (-0.454%) below 49% of LTAR. For the full sample, these results show significant negative impacts on the agricultural sector for cereals, well before reaching national-level threshold values used for defining a drought event by India's Government.

Table 6.2: Main results - Full sample and AEZs

	Full Sample	Arid	Semi-arid	Sub-humid	Humid
			Threshold test		
P-value					
Single	0	0.023	0	0	0.145
Double	0	0.129	0	0	0.64
Triple	0.602	0.631	0.51	0.622	0.698
			Threshold Location		
γ_1	0.492	0.791	0.587	0.763	
	[0.481,0.512]	[0.771,0.793]	[0.580,0.596]	[0.756,0.767]	
γ_2	0.743		0.891	0.939	
	[0.740,0.752]		[0.883,0.893]	[0.933,0.941]	
γ_3					
			β - Rainfall-Temperature Index		
Rain$< \gamma_1$	-0.606***	-0.599***	-0.521***	-0.224***	-0.194***
	(0.074)	(0.170)	(0.060)	(0.029)	(0.064)
$\gamma_1 <$ Rain $< \gamma_2$	-0.309***	0.206	-0.136***	-0.060**	
	(0.025)	(0.196)	(0.034)	(0.026)	
$\gamma_2 <$ Rain $< \gamma_3$	-0.048**		0.101**	0.084***	
	(0.023)		(0.044)	(0.029)	
Rain $> \gamma_3$					
Constant	0.052***	-0.13	0.048**	0.069***	0.065
	(0.019)	(0.098)	(0.024)	(0.021)	(0.045)
District fixed effects	✓	✓	✓	✓	✓
Year fixed effects	✓	✓	✓	✓	✓
District-specific trends					
Controls					
Grid	300	300	300	300	
Observations	12100	1012	5412	4884	792
N districts	275	23	123	111	18
R-squared	0.131	0.27	0.174	0.201	0.108

*, **, *** denote statistical significance at the 10%, 5% and 1% level, respectively. Numbers in parentheses represent clustered standard errors at the district level. All numbers in the table were rounded to 3 decimal places. This specifications uses a Balanced sample, which is a requirement for the threshold regression. The dependent variable is de-trended cereal yield where a district-specific quadratic trend was used to de-trend the variable.

Table 6.3: Summary of coefficients (all specifications)

Sample	β	T1	β	T2	β	T3	β
Full sample	-0.606***	**0.492**	-0.309***	**0.743**	-0.048**		
%Δ y for 0.01 ↑ in RTI [1]	*-0.454*		*-0.266*		*-0.047*		
Agro-ecological zones							
Arid	-0.599***	**0.791**	0.206				
%Δ y for 0.01 ↑ in RTI	*-0.451*		*0.229*				
Semi-arid	-0.521***	**0.587**	-0.136***	**0.891**	0.101**		
%Δ y for 0.01 ↑ in RTI	*-0.406*		*-0.127*		*0.106*		
Sub-humid	-0.224***	**0.763**	-0.060**	**0.939**	0.084***		
%Δ y for 0.01 ↑ in RTI	*-0.201*		*-0.058*		*0.088*		
Humid	-0.194***	**N.A.**					
%Δ y for 0.01 ↑ in RTI	*-0.176*						
Crops							
Barley	-0.136***	**N.A.**					
%Δ y for 0.01 ↑ in RTI	*-0.127*						
Maize	-0.359***	**0.576**	0.088**	**0.874**	0.269***		
%Δ y for 0.01 ↑ in RTI	*-0.302*		*0.092*		*0.309*		
Millet	-0.929***	**0.482**	-0.430***	**0.62**	0.01		
%Δ y for 0.01 ↑ in RTI	*-0.605*		*-0.349*		*0.010*		
Rice	-0.614***	**0.588**	-0.291***	**0.882**	0.008		
%Δ y for 0.01 ↑ in RTI	*-0.459*		*-0.252*		*0.008*		
Sorghum	-0.514***	**0.587**	0.009	**0.819**	0.303***		
%Δ y for 0.01 ↑ in RTI	*-0.402*		*0.009*		*0.354*		
Wheat	-0.247***	**0.798**	-0.112***	**0.955**	0.018		
%Δ y for 0.01 ↑ in RTI	*-0.219*		*-0.106*		*0.018*		

*, **, *** denote statistical significance at the 10%, 5% and 1% level, respectively. Numbers in bold denote the estimated threshold values of proportion of rain against LTAR. T1, T2 and T3 denote threshold 1, 2 and 3, respectively. Numbers in italics represent the predicted effect of a 0.01 increase in the RTI. Since the y variable is de-trended, the coefficients represent impacts in terms of deviations from trend. For example, for a coefficient of 0.136 means that, for a given event in a semi-arid area where the proportion of rain was between 58.7-89.1% of LTAR, a 0.01 increase in the index leads to a negative deviation of yield from trend of 0.127 percent.

[1] marginal effects were calculated using the formula $100 * (e^{beta} - 1)$ which gives us the marginal effect for a one unit increase the RTI. However, since a one unit increase in the RTI does not make much sense (it is the maximum value), we divide the result from the formula by 100.

Agro-ecological zone

Results by agro-ecological zone can also be seen in Table 6.2. We note that thresholds and impacts vary substantially. In arid areas, for values of rainfall above 79% of LTAR, an increase in our index (i.e. less water and/or higher temperature) leads to a statistically insignificant increase in cereal productivity, suggesting that moderate, rather than extreme excessive rainfall is preferable. Below 79% of LTAR, however, increases in our index lead to large losses in productivity, with a 0.01 increase in the index associated with a negative deviation from trend of -0.599%. We note that negative significant impacts of increases in the index (i.e. decreases in rainfall/increases in temperature) appear to start at around 79% of LTAR. The coefficient of the index is larger than in the full sample and the associated impacts are larger in magnitude. A potential reason for a lower threshold than that found in the full sample may be the possibility of adaptation, notably through crop choice.

In semi-arid areas, for levels of rainfall above 89% of the LTAR, an increase in the index is found to have a positive marginal impact (+0.106% from trend for a 0.01 increase in index), which suggests higher yields at more moderate levels of rainfall (closer to the LTAR). The threshold, at which impacts become negative, is higher than in arid areas (below 89% of LTAR) but between 89%-59% of LTAR, impacts are high (0.127%), though smaller than for arid areas for comparable levels of rainfall deficiency. Below 59% of LTAR, the marginal effects of decreased rainfall/increased temperature are likely to have extremely severe impacts on agricultural productivity (-0.406% from trend).

In sub-humid, our identified thresholds tend to be higher. The threshold in the former is estimated at a similar level to semi-arid areas (94% of LTAR). We find a positive coefficient for the index for levels above 94% of LTAR. This suggests higher yields at normal rainfall compared to extreme positive rainfall, although no flood threshold is identified. Negative impacts from the increase in our index start below 94% of LTAR and remain moderate (-0.06% from trend for a 0.01 increase in the index) until rainfall below 76% of LTAR (-0.224%). We note that the impacts tend to be substantially worse than for similar levels of rainfall deficiencies in more arid areas.

Finally, in humid areas, the existence of a threshold is not supported at the conventional statistical levels. As a result, the coefficient shown refers to a standard fixed effects regression. The coefficient implies that the marginal effects tend to be quite small (-0.176% from trend for a 0.01 increase in the index) for large negative deviations in rainfall compared with other agro-ecological zones, although at low deviations from LTAR impacts are high. One explanation could be the fact that these areas are less likely to be irrigated and tend to cultivate more water-intensive crops. As a result, small deviations in rainfall may have a negative impact on their productivity. Nevertheless, as these areas typically have abundant rainfall, this also means that crop water stress is unlikely to be as severe as in more arid areas.

Crops

Different crops are known for different water requirements and resistance to water stress. The results by crop are presented in Table 6.4.

Table 6.4: Main results - Crops

	Barley	Maize	Millet	Rice	Sorghum	Wheat
			Threshold test - P-value			
Single	0.837	0	0	0	0	0
Double	0.348	0	0	0	0	0.001
Triple	0.719	0.545	0.575	0.448	0.59	0.434
			Threshold Location			
γ_1		0.576	0.482	0.588	0.587	0.798
		[0.566,0.583]	[0.466,0.502]	[0.577,0.618]	[0.580,0.598]	[0.787,0.800]
γ_2		0.874	0.62	0.882	0.819	0.955
		[0.855,0.876]	[0.616,0.625]	[0.877,0.884]	[0.814,0.821]	[0.951,0.958]
γ_3						
			β - Rainfall-Temperature Index			
Rain$< \gamma_1$	-0.136***	-0.359***	-0.929***	-0.614***	-0.514***	-0.247***
	(0.041)	(0.070)	(0.130)	(0.066)	(0.076)	(0.029)
$\gamma_1 <$ Rain $< \gamma_2$		0.088**	-0.430***	-0.291***	0.009	-0.112***
		(0.041)	(0.070)	(0.025)	(0.044)	(0.033)
$\gamma_2 <$ Rain $< \gamma_3$		0.269***	0.01	0.008	0.303***	0.018
		(0.053)	(0.038)	(0.032)	(0.056)	(0.040)
Rain $> \gamma_3$						
Constant	0.034	-0.167***	-0.024	0.150***	-0.045	0.186***
	(0.032)	(0.039)	(0.034)	(0.020)	(0.047)	(0.019)
District fixed effects	✓	✓	✓	✓	✓	✓
Year fixed effects	✓	✓	✓	✓	✓	✓
District-specific trends						
Controls						
Grid		300	300	300	300	300
Observations	3432	7656	7172	10560	6908	8756
N districts	78	175	163	240	157	199
R-squared	0.16	0.106	0.144	0.165	0.106	0.088

*, **, *** denote statistical significance at the 10%, 5% and 1% level, respectively. Numbers in parentheses represent clustered standard errors at the district level. All numbers in the table were rounded to 3 decimal places. This specifications uses a Balanced sample, which is a requirement for the threshold regression. The dependent variable is de-trended yield for each individual cereal, where a district-specific quadratic trend was used to de-trend the variable.

Our results suggest that rice is the most sensitive crop to negative rainfall deviations. Our model supports the existence of two thresholds (both below the LTAR), and suggests an insignificant positive coefficient for our index for levels of rainfall above 88% of LTAR (+0.008%

from trend for a 0.01 increase in the index). However, even at small deviations from LTAR (below 88% of LTAR), the index exhibits a large negative coefficient (-0.252% from trend). This coefficient becomes substantially more negative at levels of rainfall below 59% of LTAR, where impacts become very large (-0.459% from trend).

For wheat, the estimated thresholds are similar to those of rice (95% and 80% of LTAR). Impacts on wheat are a lot smaller for similar rainfall deficiencies with estimated impacts at levels of rainfall below 95% and 80% LTAR estimated at -0.106% and -0.219% for a 0.01 increase in the index, respectively. There are two potential explanations for this. Looking at Figures 1 and 2, we can see that rice tends to be grown mostly in humid and sub-humid areas, whereas wheat tends to be more concentrated in semi-arid areas. Such areas have very different agro-ecological conditions. Also, the proportion of area irrigated differs, which corroborates with Singh et al. (2011), who document that a much higher proportion of wheat is cultivated under irrigation in comparison to rice.

In the case of barley, the identified thresholds are not statistically significant and thus we estimate the regression using a standard fixed effects model. The estimated coefficient suggests an impact of moderate magnitude (-0.127% from trend for a 0.01 increase in the index).

The remaining crops (maize, sorghum and millet) are often considered to be more drought-resistant than rice, wheat and barley and this is corroborated by our results. For maize, we estimate two thresholds, both below the LTAR. These suggest that negative impacts are felt at very large negative deviations from the LTAR (below 58% of LTAR). The estimated impacts are very large (a deviation from trend of -0.302% is estimated for a 0.01 increase in the index). For both sorghum and millet, we also find two thresholds, all below the LTAR. The results for millet suggest that for levels of rainfall below 62% of LTAR, there are large negative impacts (-0.349% for a 0.01 increase in the index), which become extremely large for levels of rainfall below 48% of LTAR (-0.605% for a 0.01 increase in the index). Finally, for sorghum we find a positive coefficient above 82% of LTAR and insignificant impacts between 82% and 59% of LTAR. However, below 59% we find very large impacts (-0.402%). It is perhaps surprising that two of the crops widely considered as being the most drought-resistant exhibit the largest impacts of decreases in rainfall at low levels of rainfall. A possible explanation is that, while these crops are more drought-resistant and therefore have a lower threshold, they are mainly grown under rain-fed conditions and hence, may be more vulnerable to increasingly erratic patterns of rainfall.

Further Results: Why so few excess rain thresholds?

The vast majority of our estimated thresholds are for levels of rainfall below average. There are at least three potential explanations for why this could be the case, even if there are large, significant impacts of excessive rainfall. First, protracted periods of low rainfall may have a more pernicious effect on agricultural productivity. Singh et al. (2011), for instance, find that only two out of eight "flood" years (defined as mean + one standard deviation) led to an aggregate loss in grain production in India. Secondly, Collier and Webb (2012) argue that floods tend to be more localized in space and time and as such we may be unable to capture these because the data are collected at the district level, and rainfall is aggregated over a whole year. Finally, our method requires us to trim some portions of the data at each extremity of the threshold variable (proportion of rainfall). Aufhammer et al. (2012) find negative impacts for rainfall in excess of the 95th percentile. Perhaps, excess rainfall only has very negative effects for very extreme rainfall deviations. If this deviation is very close to the last percentiles, our method may not capture this since the threshold may only be present in the trimmed portion.

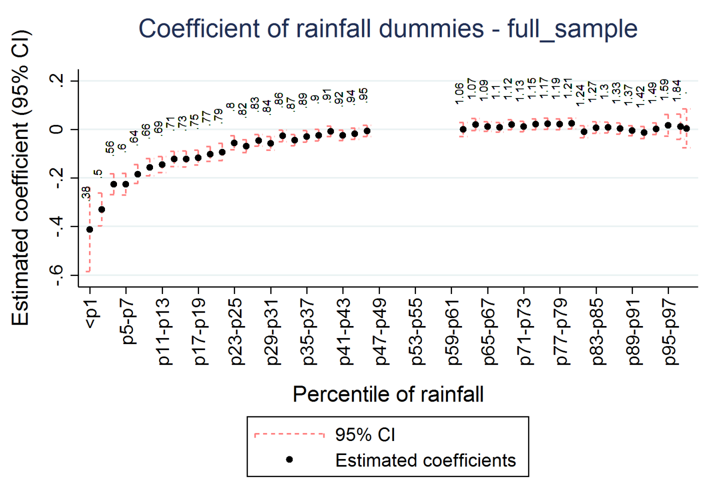
As a robustness check, we run a regression of yield with dummies highlighting different percentiles of rainfall[15]. This approach is complementary to the one used in the previous subsection as it shows at which points impacts are likely to become significant rather than at what proportion of rainfall the marginal effects are likely to change. We create dummies for every second percentile of rainfall against the LTAR, where dummy is compared to the baseline category, set at LTAR +/-5%. The main results from this approach are summarized in Figures 6.4-6.6 (by agro-ecological zone) and 6.7-6.9 (by crop). For the full sample [figure 6.4, panel (a)], we find no effect of consistent negative significant effects of excessive rainfall on cereal productivity. However, for semi-arid areas [Figure 6.5, panel (b)] and, to some extent, for semi-humid areas [Figure 6.6, panel (a)] we find significant, but relatively modest negative

[15]One of the reasons we do not adopt this method as the main method is that it is extremely difficult to incorporate both temperature and rainfall non-parametrically in this way, due to the huge number of interactions required. While there also seems to be no clear theory regarding how excess rainfall interacts with temperature, this is a useful method for the purpose of understanding the impacts from excess rainfall.

effects of excess rainfall but only for very extreme events. These effects occur beyond the 89^{th} percentile in semi-arid areas, becoming large above the 99^{th} percentile[16]. For sub-humid areas, the negative effects are only found to be significant for years where rainfall exceeds the 93^{rd} and the 99^{th} percentile of rainfall[17]. We find consistent positive impacts for high levels of rainfall in arid areas, which suggests that optimal rainfall ranges may be above the LTAR. In humid areas, the estimation is noisier and it is hard to discern any pattern. This could be due to sample size.

Estimates by crop also highlight negative effects of excessive rainfall for maize, sorghum and, to a lesser extent, millet. The effects on maize are very large (comparable or even higher than for drought) and start beyond the 81^{st} percentile (123% of LTAR). For sorghum negative impacts from excessive rainfall also begin around the 81^{st} percentile and become significantly larger beyond the 95^{th} percentile. Finally, in the case of millet, impacts are smaller and start later, around the 95^{th} percentile.

Figure 6.4: Non-Parametric results (rainfall dummies) full sample

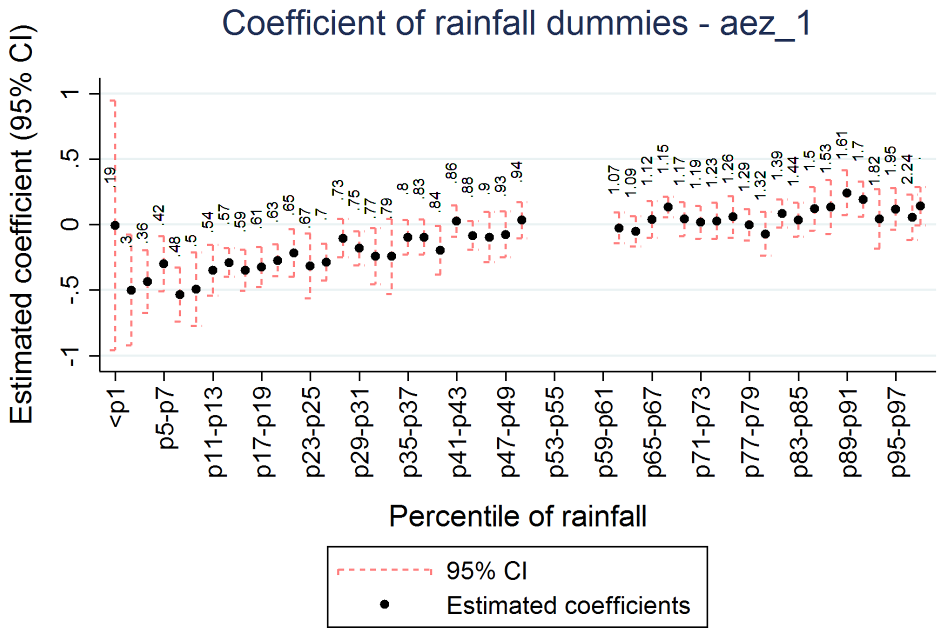


(a) Full sample

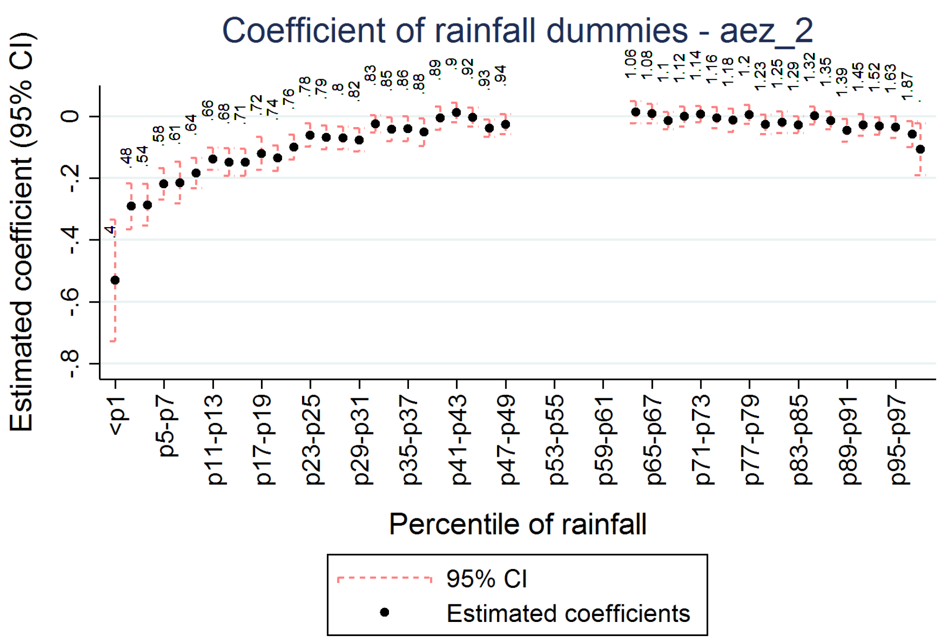
Confidence bands indicate 95% confidence intervals. The numbers above the scatter points indicate the maximum proportion of rain (vs. the LTAR) included in the estimated dummy variable. Finally, in the x-axis p denotes the percentiles of rainfall. As a result, p5-p7 should be interpreted as rainfall (as a proportion of the LTAR) between the fifth and the seventh percentile.

[16]In semi-arid areas the 89^{th} percentile corresponds roughly to 139% of LTAR.
[17]This (93rd percentile) corresponds to rainfall in excess of 142% of LTAR.

Figure 6.5: Non-Parametric results (rainfall dummies) by agro-ecological zone - Arid areas



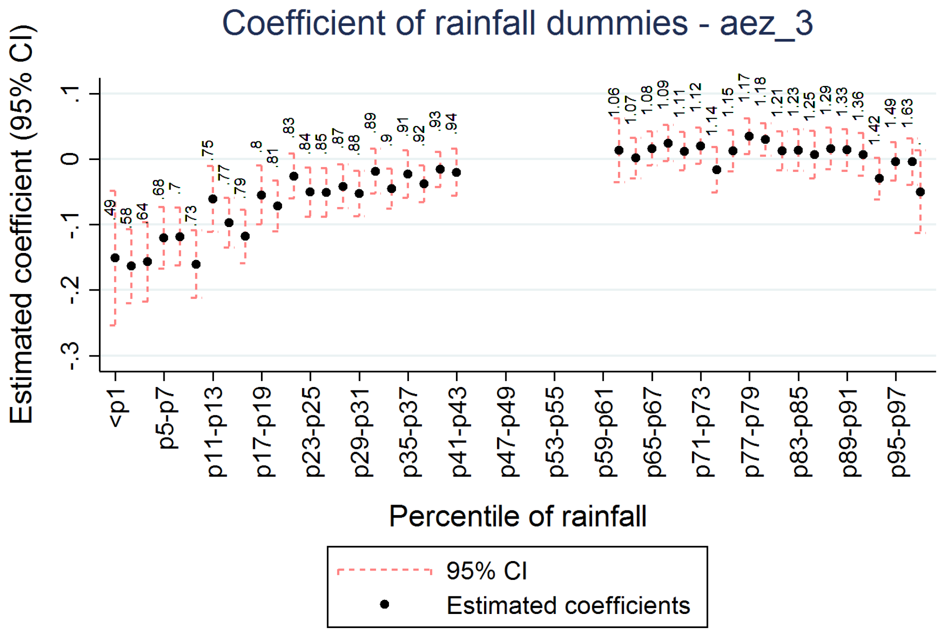
(a) Arid areas



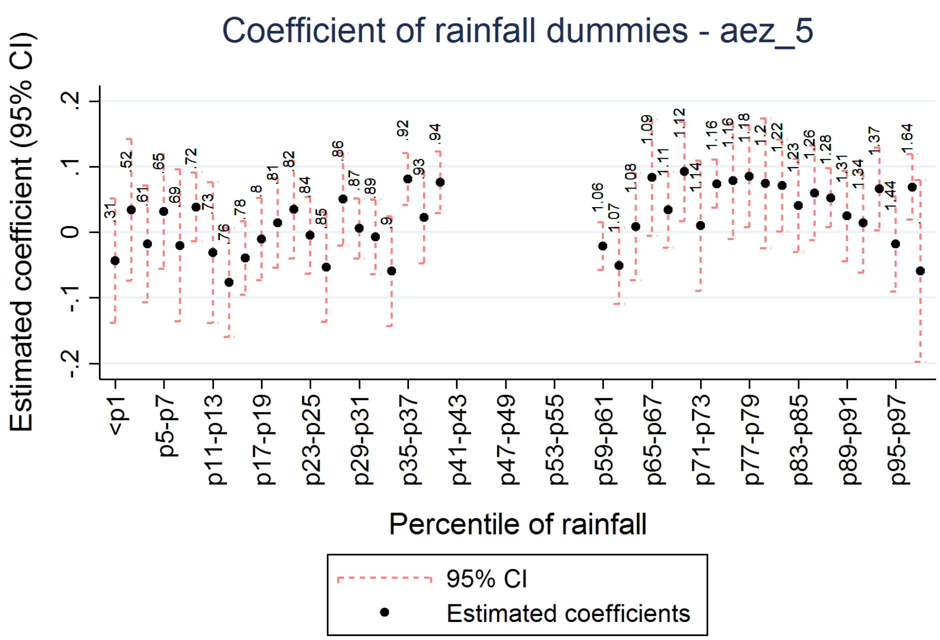
(b) Semi-arid areas

Confidence bands indicate 95% confidence intervals. The numbers above the scatter points indicate the maximum proportion of rain (vs. the LTAR) included in the estimated dummy variable. Finally, in the x-axis p denotes the percentiles of rainfall. As a result, p5-p7 should be interpreted as rainfall (as a proportion of the LTAR) between the fifth and the seventh percentile.

Figure 6.6: Non-Parametric results (rainfall dummies) by agro-ecological zone - humid areas



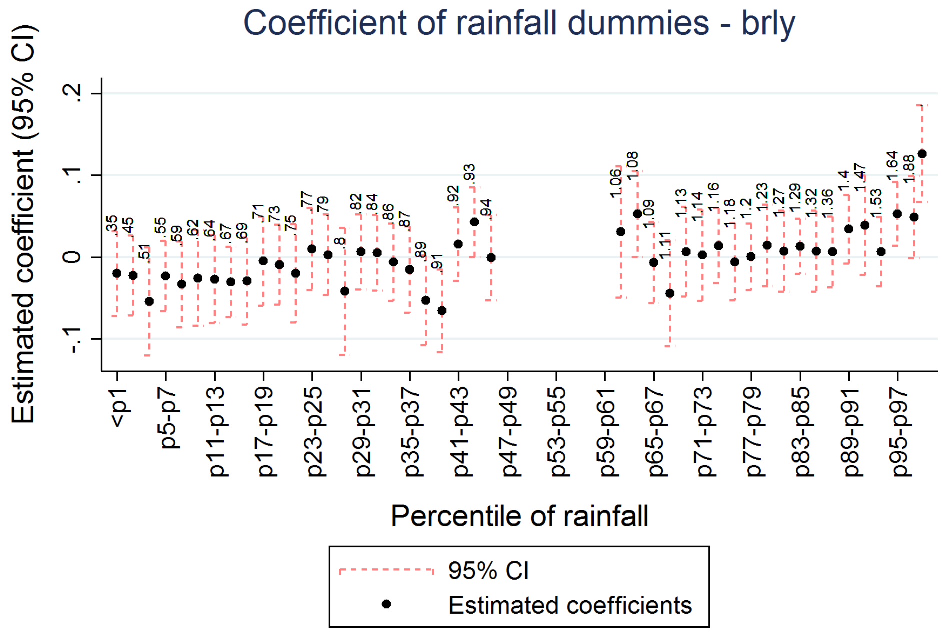
(a) Sub-humid areas



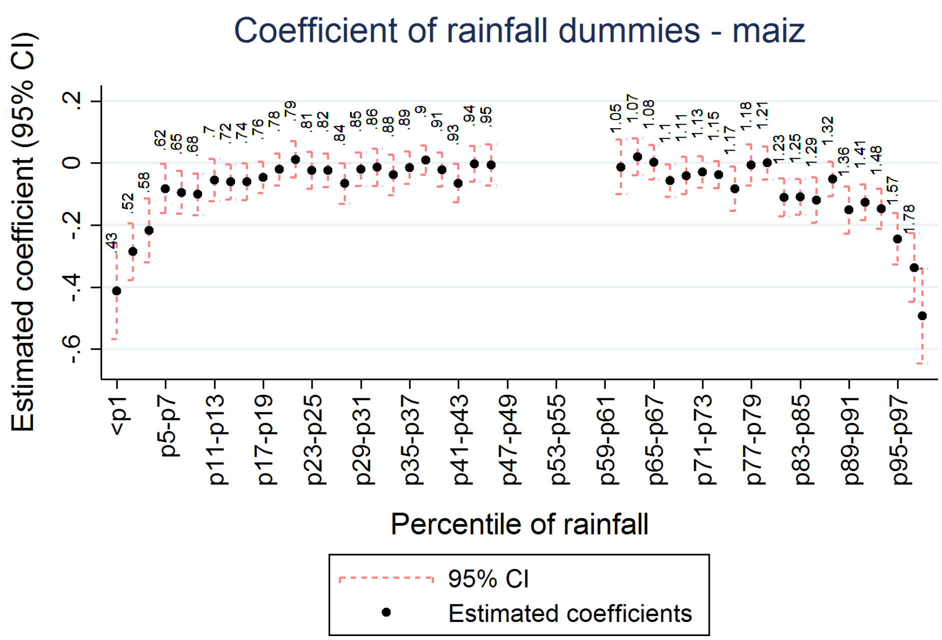
(b) Humid areas

Confidence bands indicate 95% confidence intervals. The numbers above the scatter points indicate the maximum proportion of rain (vs. the LTAR) included in the estimated dummy variable. Finally, in the x-axis p denotes the percentiles of rainfall. As a result, p5-p7 should be interpreted as rainfall (as a proportion of the LTAR) between the fifth and the seventh percentile.

Figure 6.7: Non-Parametric results (rainfall dummies) by crop - barley and maize



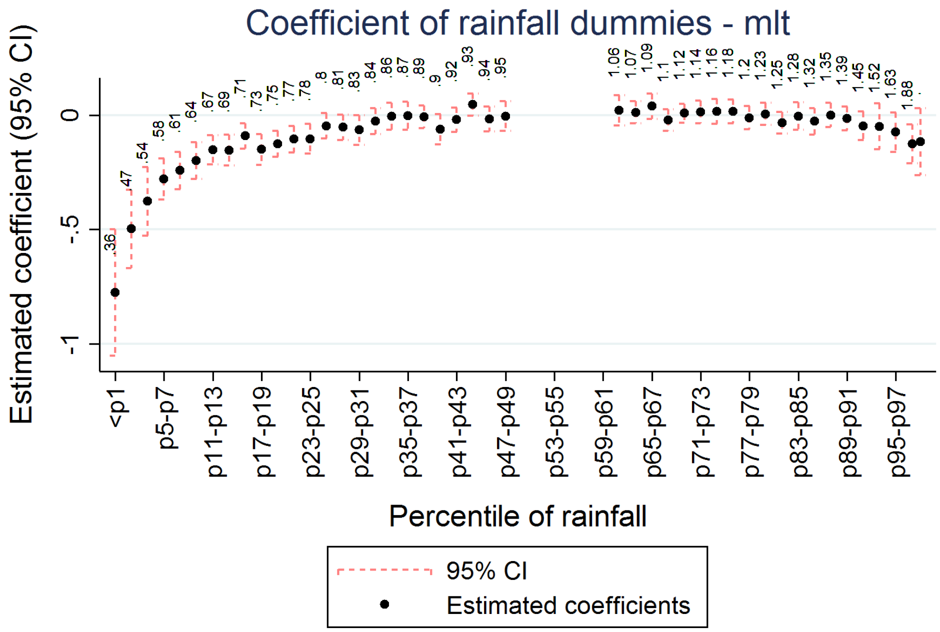
(a) Barley



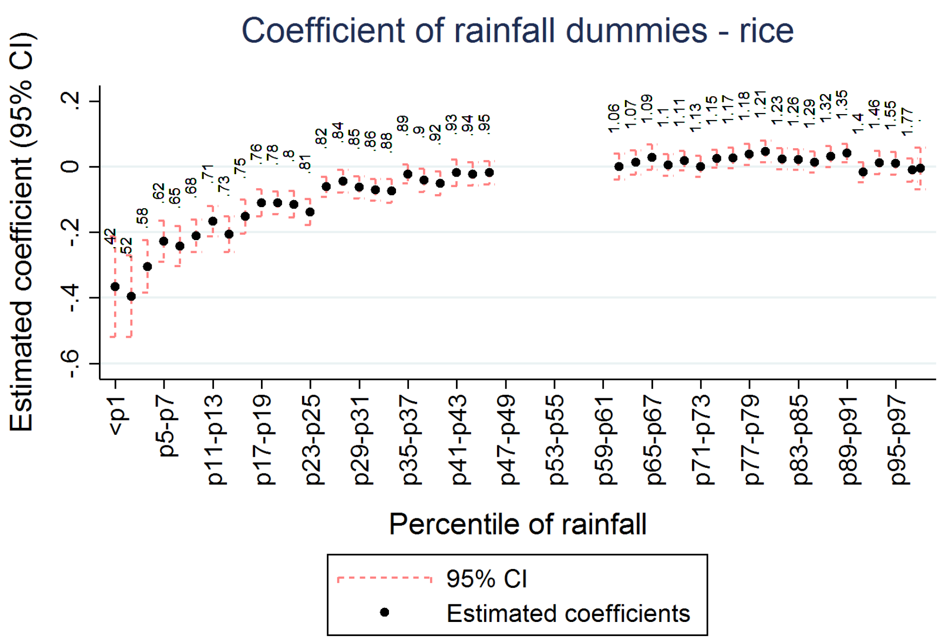
(b) Maize

Confidence bands indicate 95% confidence intervals. The numbers above the scatter points indicate the maximum proportion of rain (vs. the LTAR) included in the estimated dummy variable. Finally, in the x-axis p denotes the percentiles of rainfall. As a result, p5-p7 should be interpreted as rainfall (as a proportion of the LTAR) between the fifth and the seventh percentile.

Figure 6.8: Non-Parametric results (rainfall dummies) by crop - millet and rice



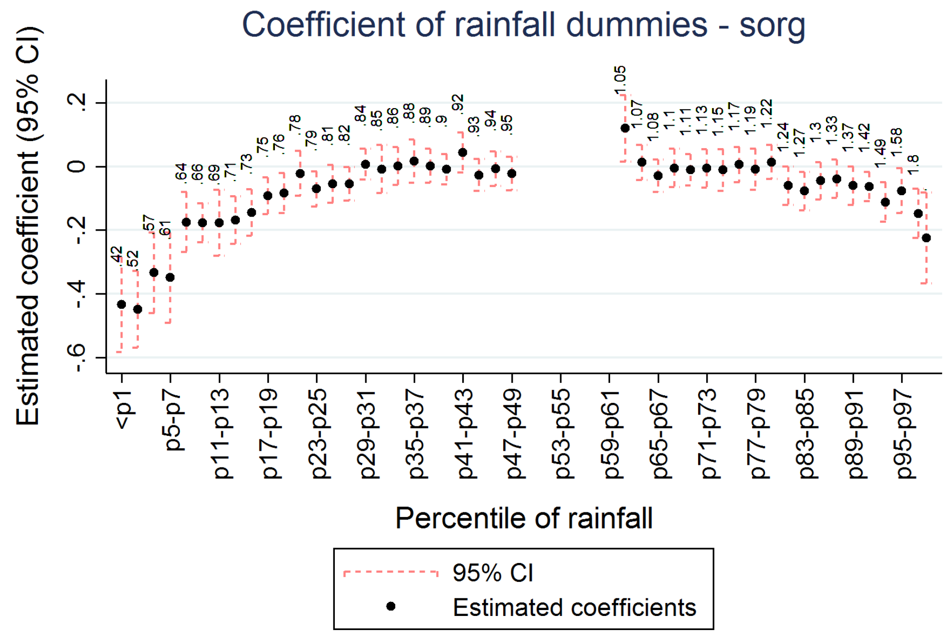
(a) Millet



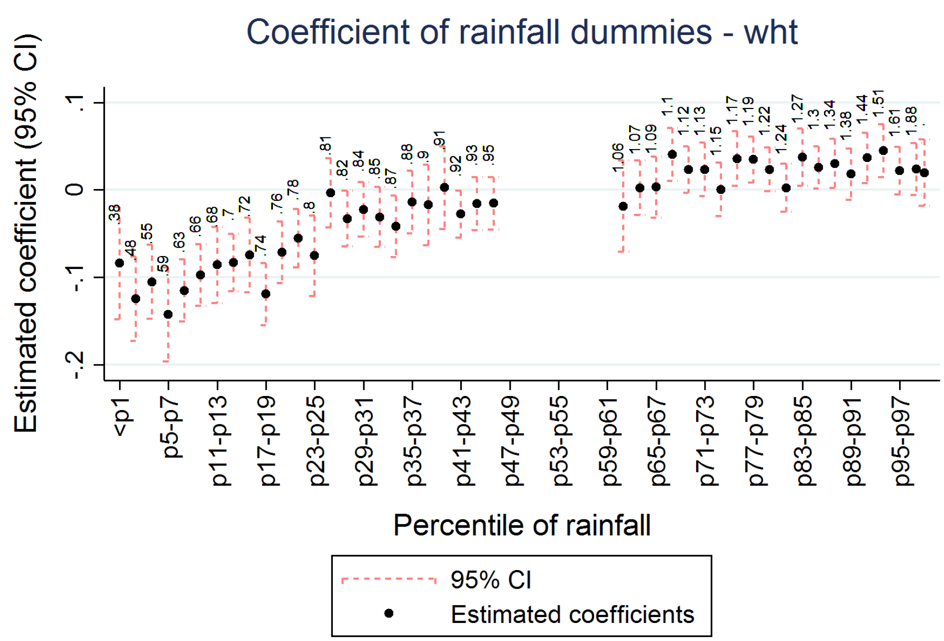
(b) Rice

Confidence bands indicate 95% confidence intervals. The numbers above the scatter points indicate the maximum proportion of rain (vs. the LTAR) included in the estimated dummy variable. Finally, in the x-axis p denotes the percentiles of rainfall. As a result, p5-p7 should be interpreted as rainfall (as a proportion of the LTAR) between the fifth and the seventh percentile.

Figure 6.9: Non-Parametric results (rainfall dummies) by crop - sorghum and wheat



(a) Sorghum



(b) Wheat

Confidence bands indicate 95% confidence intervals. The numbers above the scatter points indicate the maximum proportion of rain (vs. the LTAR) included in the estimated dummy variable. Finally, in the x-axis p denotes the percentiles of rainfall. As a result, p5-p7 should be interpreted as rainfall (as a proportion of the LTAR) between the fifth and the seventh percentile.

In sum, our results suggest that there are significant negative effects of excess rainfall on crop productivity but that these are not picked up by the threshold model. We argue that this is likely due to the trimming. Most impacts occur at very extreme positive deviations in rainfall and, as such, this threshold is either 1) likely not even to be considered, 2) the estimation is likely to be noisy given the small sample size on the positive side (excess rain) of the threshold. In some cases (e.g. maize) the threshold above the LTAR is identified, but it is not statistically significant.

In addition to this, this estimation also allows us to confer whether our threshold estimation results are plausible. For the full sample and the agro-ecological zones, the thresholds seem very plausible and consistent with the non-parametric analysis and the same applies to the crop estimates for rice, wheat and maize. However, in the case of millet and sorghum we notice that there are negative impacts before the thresholds we estimated. In addition to this, it is important to remember that the threshold regression estimates the points at which there are likely to be changes in the coefficients and the identified thresholds for both millet and sorghum often occur at the points where we notice an abrupt increase in impacts. In addition, the threshold regression identified a third threshold for millet (at 79% of LTAR) and for sorghum (at 72% of LTAR). However, these were not statistically significant.

Robustness checks

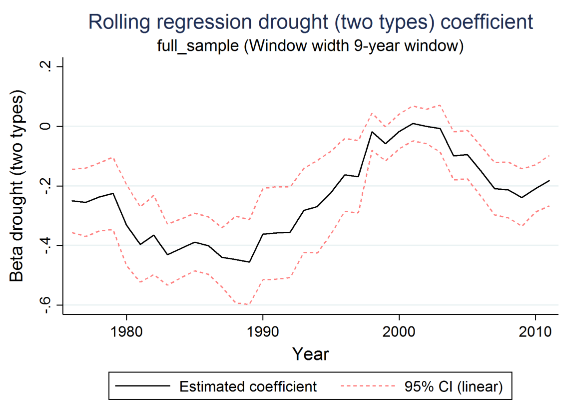
We analysed the sensitivity of our results to a number of assumptions and a number of specifications. Our preferred specification is a reduced-form function where controls are excluded. This is arguably the most common type of specification in the climate literature. However, given that our method requires a balanced panel, we also lose a number of districts. In the case of the full sample, we lose approximately 73 (out of 275) districts by including only 4 controls. To ensure that our results are not affected by the omission of control variables, we include controls in Tables 6A.3 and 6A.4 (see Appendix). In addition to this, we also tested the robustness of our results to: 1) an alternative dependent variable (de-meaned, rather than de-trended yield) and results can be found in tables 6A.5 and 6A.6; 2) alternative growing seasons (Annual and May-December, Tables 6A.7-6A.10); and 3) alternative index (additive relationship) (tables 6A.11 and 6A.12). In most cases, the robustness checks provide qualitative results that are similar to those of the main results.

6.4.2 Impacts of drought over time

Shape of the impacts over time

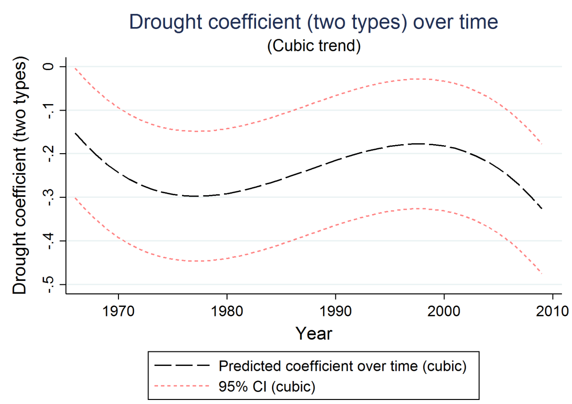
We begin by estimating the rolling regression for the full sample (Figure 6.10) and the coefficient plot for the 9-year window reveals what appears to be a non-linear pattern. According to the rolling regression estimates, impacts of drought became smaller between the late seventies and the late nineties, following which there seems to be a reversal in the trend. The non-linear pattern is confirmed when we estimate our preferred parametric model (as suggested by the F-test in Table 6A.13) and a cubic trend is preferred using the test of joint significance and the shape reproduced by the parametric model is not very different from that of the rolling regression.

Figure 6.10: Time results - Full sample



(a) Full sample - rolling regression

(9-year window)



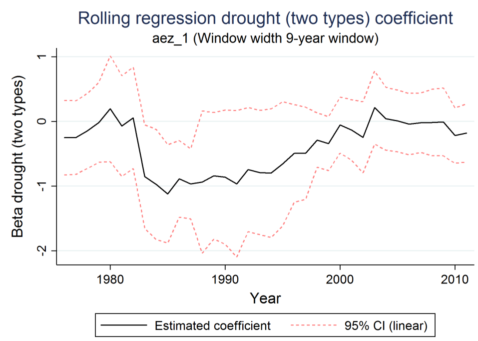
(b) Full sample - parametric fit

(cubic interaction)

Rollling regressions with different windows are shown in Appendix A. In order to choose the parametric fit we ran three regressions (linear, quadratic and cubic interaction between the RTI and the time trend). The parametric fit chosen was the specification for which the joint significance test had the lowest p-value. The F- and p-values can be seen in Table 6A.13 in Appendix A. In some occasions, the difference in p-values was very small between specifications. In these cases we report the preferred specification and include the figures of the other specifications in the Appendix. Finally, in other cases, none of the F-tests was significant. In these cases, we only report the rolling regressions and the coefficient tables can be found in the Appendix.

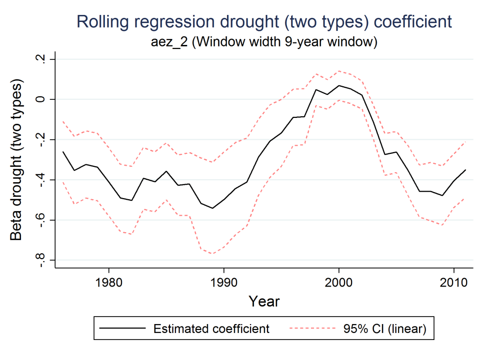
A similar non-linear pattern is also found for semi-arid and sub-humid areas (Figures 6.11 and 6.12) and for both a cubic trend is preferred. For humid areas and arid areas, however, either there is no support for a trend (arid) or a linear trend is preferred (humid), though it is unclear whether this is due to the small sample sizes of these sub-samples. Most strikingly, we observe a sharp decline in drought impacts starting from the eighties and then a sharp increase in the impacts starting from the late nineties/early 2000s.

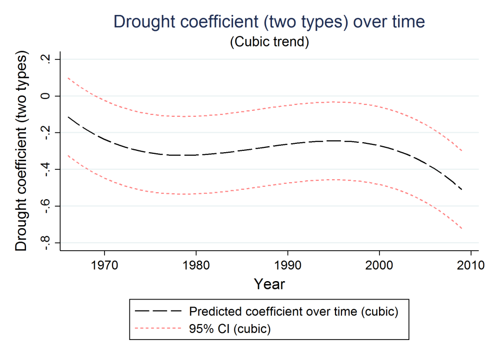
Figure 6.11: Time results - arid and semi-arid areas



(a) Arid areas - rolling regression

(9-year window)





(b) Semi-arid areas - rolling regression

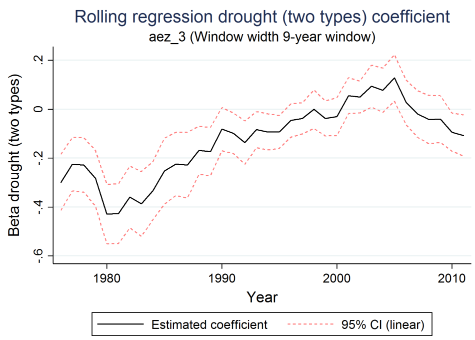
(9-year window)

(c) Semi-arid areas - parametric fit

(cubic interaction)

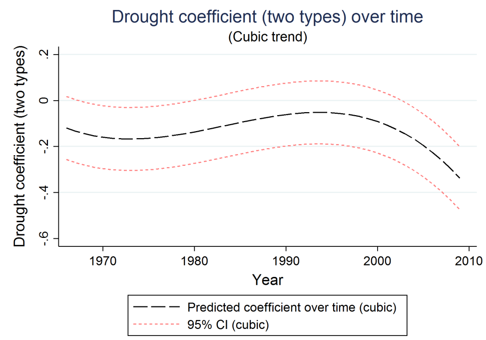
Rollling regressions with different windows are shown in Appendix A. In order to choose the parametric fit we ran three regressions (linear, quadratic and cubic interaction between the RTI and the time trend). The parametric fit chosen was the specification for which the joint significance test had the lowest p-value. The F- and p-values can be seen in Table 6A.13 in Appendix A. In some occasions, the difference in p-values was very small between specifications. In these cases we report the preferred specification and include the figures of the other specifications in the Appendix. Finally, in other cases, none of the F-tests was significant. In these cases, we only report the rolling regressions and the coefficient tables can be found in the Appendix. This is why the parametric fit for arid areas is not reported.

Figure 6.12: Time results - sub-humid and humid areas



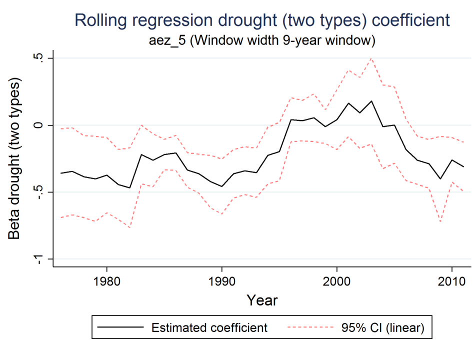
(a) Sub-humid areas - rolling regression

(9-year window)



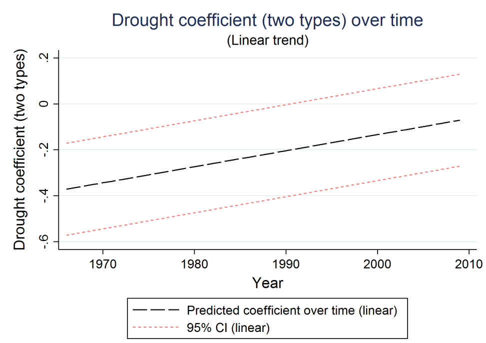
(b) Sub-humid areas - parametric fit

(cubic interaction)



(c) Humid areas - rolling regression

(9-year window)



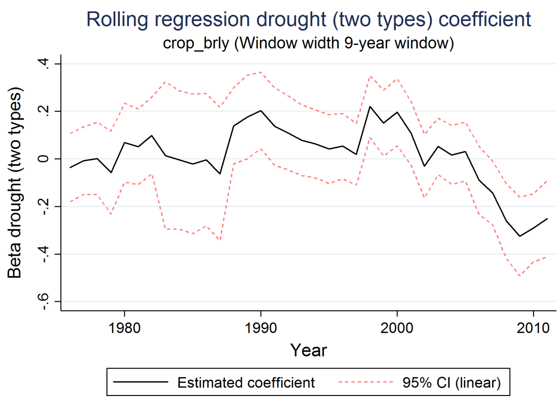
(c) Semi-arid areas - parametric fit

(linear interaction)

Rollling regressions with different windows are shown in Appendix A. In order to choose the parametric fit we ran three regressions (linear, quadratic and cubic interaction between the RTI and the time trend). The parametric fit chosen was the specification for which the joint significance test had the lowest p-value. The F- and p-values can be seen in Table 6A.13 in Appendix A. In some occasions, the difference in p-values was very small between specifications. In these cases we report the preferred specification and include the figures of the other specifications in the Appendix. Finally, in other cases, none of the F-tests was significant. In these cases, we only report the rolling regressions and the coefficient tables can be found in the Appendix.

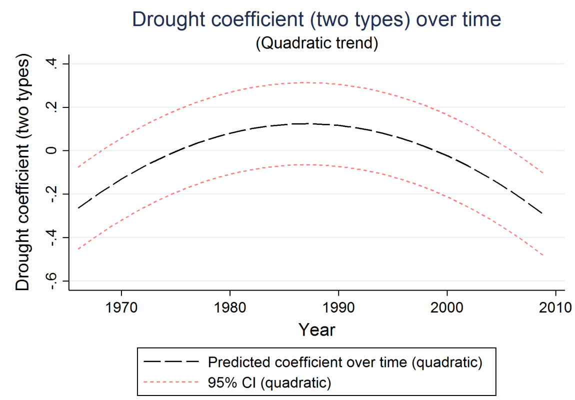
Finally, the results by crop (Figures 6.13 and 6.14) also suggest, for the most part, a non-linear pattern and the sharp drop after the 2000s is noticeable in most crops, though slightly less pronounced for rice and wheat. However, we also notice sharp differences across the shapes of the impacts over time. For instance, drought impacts of rice and wheat tend to decrease in a smoother way over time. Conversely, in the case of maize, millet and sorghum, the coefficient

Figure 6.13: Time results - barley, rice and wheat



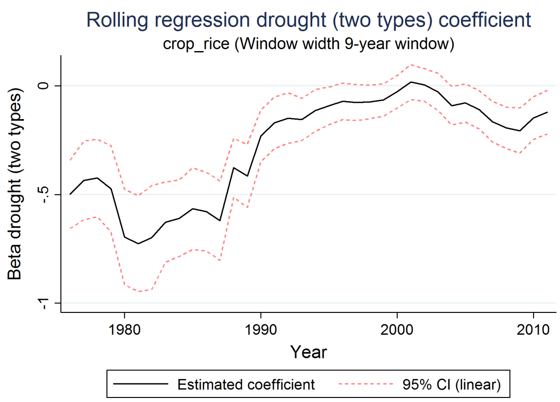
(a) Barley - rolling regression

(9-year window)



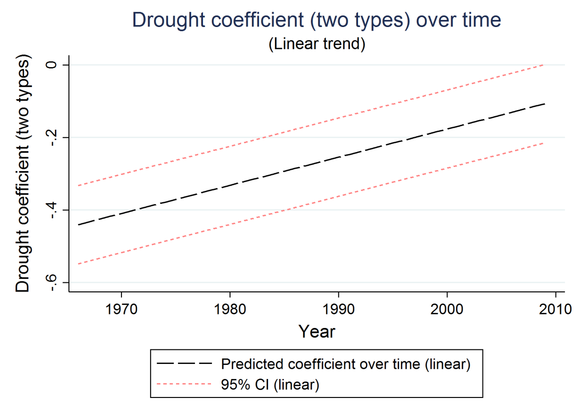
(b) Barley - parametric fit

(quadratic interaction)



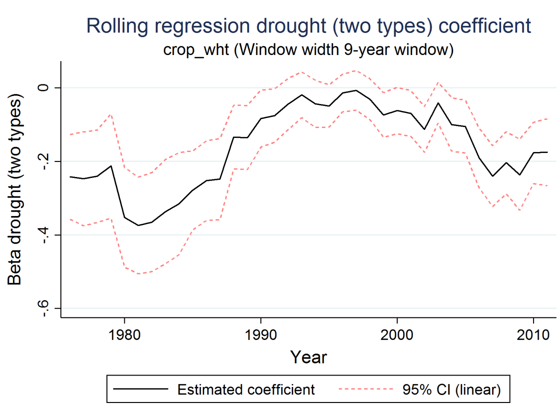
(c) Rice - rolling regression

(9-year window)



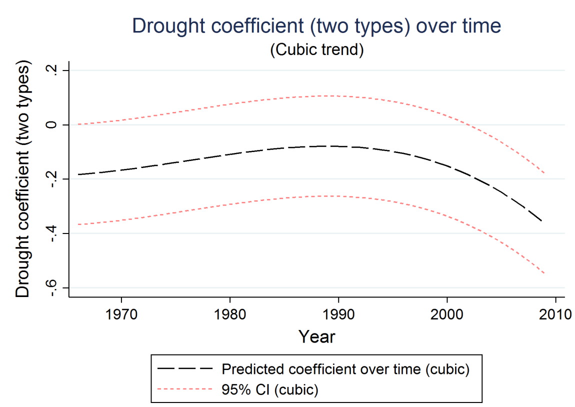
(d) Rice - parametric fit

(linear interaction)



(e) Wheat - rolling regression

(9-year window)

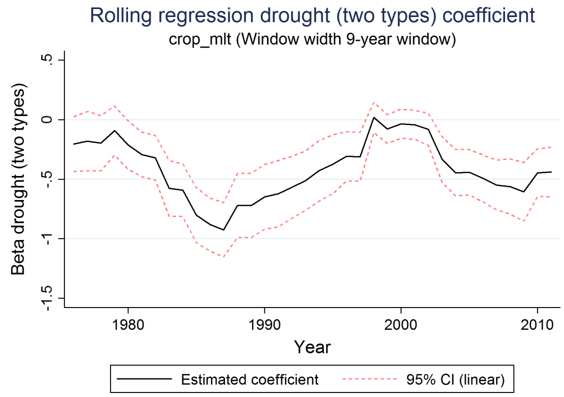


(f) Wheat - parametric fit

(cubic interaction)

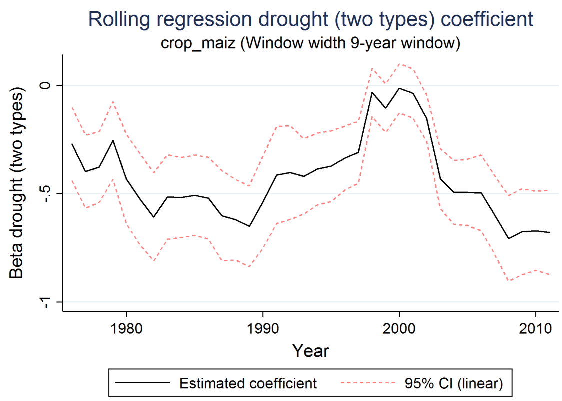
Rollling regressions with different windows are shown in Appendix A. In order to choose the parametric fit we ran three regressions (linear, quadratic and cubic interaction between the RTI and the time trend). The parametric fit chosen was the specification for which the joint significance test had the lowest p-value. The F- and p-values can be seen in Table 6A.14 in Appendix A. In some occasions, the difference in p-values was very small between specifications. In these cases we report the preferred specification and include the figures of the other specifications in the Appendix. Finally, in other cases, none of the F-tests was significant. In these cases, we only report the rolling regressions and the coefficient tables can be found in the Appendix.

Figure 6.14: Time results - maize, millet and sorghum



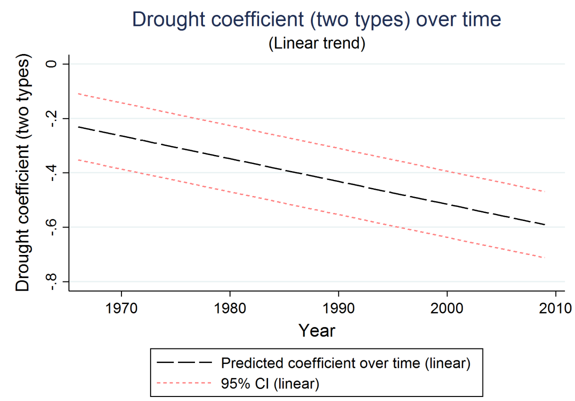
(a) Millet - rolling regression

(9-year window)



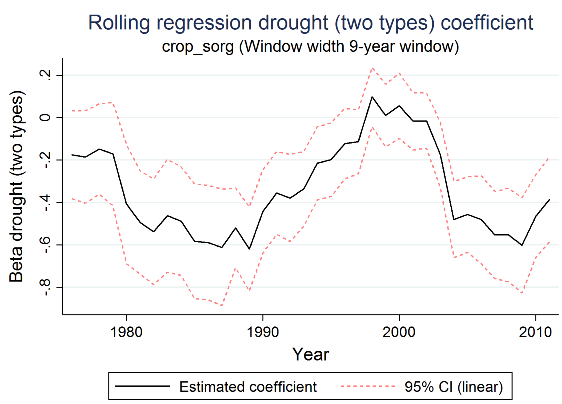
(b) Maize - rolling regression

(9-year window)



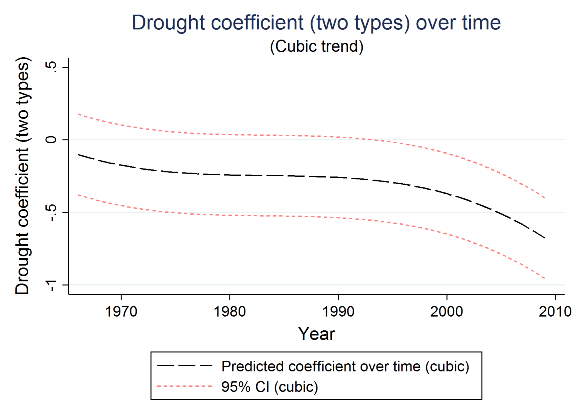
(c) Maize - parametric fit

(Linear interaction)



(e) Sorghum - rolling regression

(9-year window)



(f) Sorghum - parametric fit

(cubic interaction)

Rollling regressions with different windows are shown in Appendix A. In order to choose the parametric fit we ran three regressions (linear, quadratic and cubic interaction between the RTI and the time trend). The parametric fit chosen was the specification for which the joint significance test had the lowest p-value. The F- and p-values can be seen in Table 6A.14 in Appendix A. In some occasions, the difference in p-values was very small between specifications. In these cases we report the preferred specification and include the figures of the other specifications in the Appendix. Finally, in other cases, none of the F-tests was significant. In these cases, we only report the rolling regressions and the coefficient tables can be found in the Appendix.

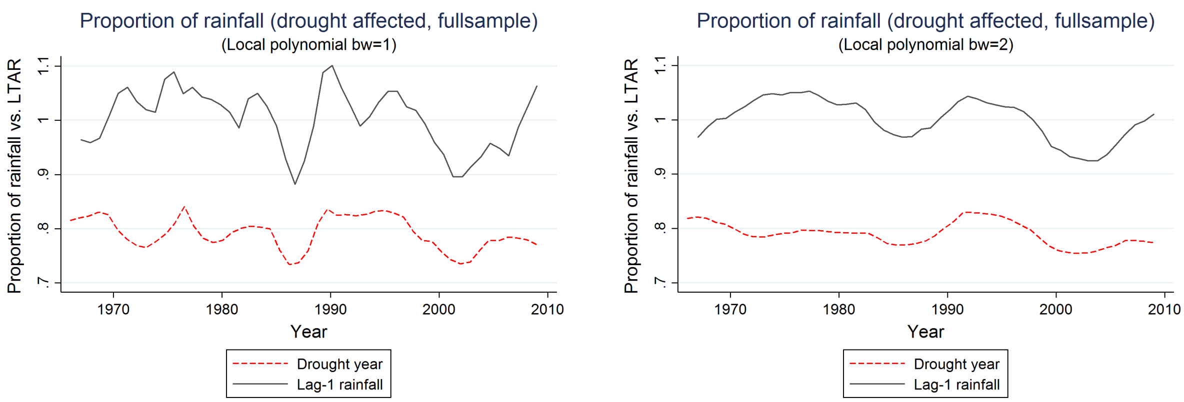
Discussion of the shape of the drought impacts over time

Three results emerge from our analysis of the impacts over time. First, the shape is non-linear in the majority of cases. Second, in most cases, we witness a decrease in the impacts until the late nineties and a reversal of the trend since. Finally, there are differences in terms of the geographical pattern and crops.

It is plausible that the decreases in impacts until the late nineties may have come as a result of the expansion of irrigated area, diffusion of improved varieties and increased use of inputs, as shown earlier (figure 6.3). However, the increase in impacts starting from the early 2000s is more puzzling. A plausible explanation for this could be attributed, at least partly, to a change in the pattern of rainfall. For instance, Shah et al. 2009 explain that in 2002/3 one of the reasons why the impacts of the drought were so high was the fact that both the 2000/1 and 2001/2 seasons had been characterized by slightly deficient rainfall. As a result, in 2002/3 when there was the drought, the aquifers had not had time to replenish, leading to dramatic impacts. This pattern, however, seems to have become more common. We run a local polynomial regression on the percentage of rainfall and the lagged precipitation received by drought-affected districts[18] and we notice that, starting from the late nineties there seems to be a drop in the lagged rainfall received by drought affected districts compared with previous periods. As shown in figure 6.15, for much of the seventies and nineties, the proportion of rainfall in the year preceding a drought was above normal and this trend seems to have been reversed in the early 2000s. We also note that this pattern is most stark in semi-arid (figure 6.16, compared to sub-humid areas (figure 6.17), which could help explain why we notice a sharper increase in the impacts in semi-arid areas.

[18] i.e. the proportion of precipitation against the LTAR received the year before the drought by drought-affected districts is smaller.

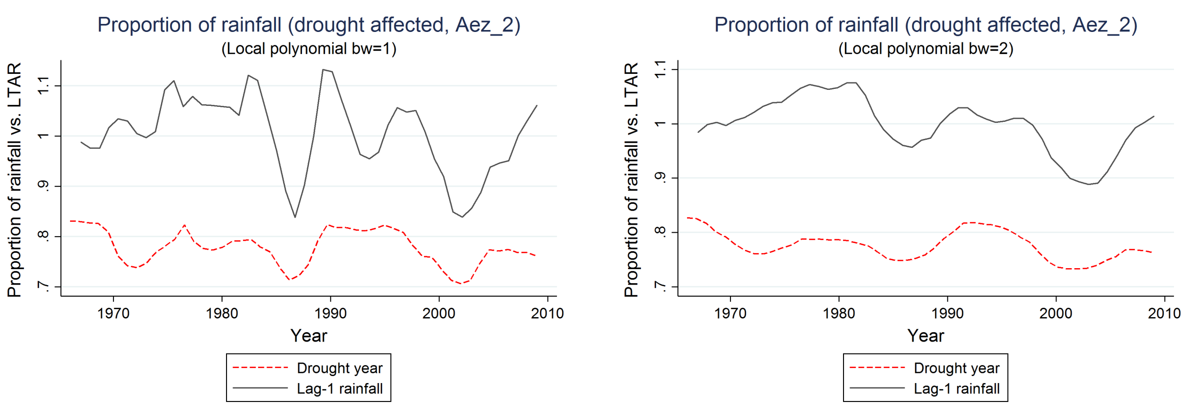
Figure 6.15: Rainfall and lagged rainfall for drought affected districts



(a) Full sample - bw=1

(b) Full sample - bw=2

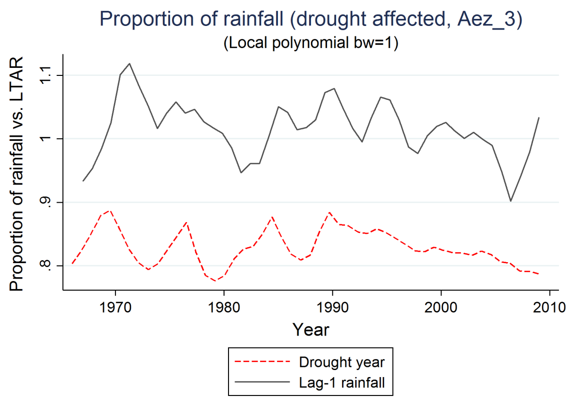
Figure 6.16: Rainfall and lagged rainfall for drought affected districts in semi-arid areas



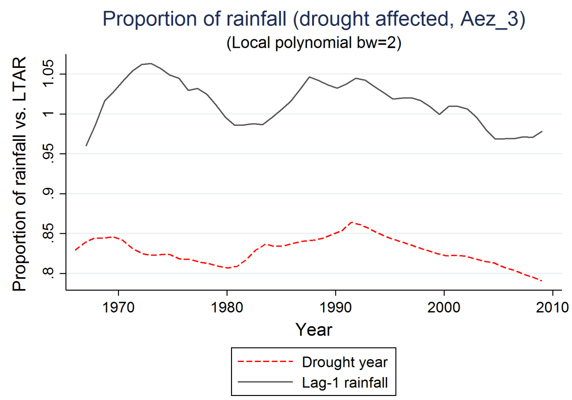
(a) Semi-arid - bw=1

(b) Semi-arid - bw=2

Figure 6.17: Rainfall and lagged rainfall for drought affected districts in sub-humid areas



(a) Sub-humid - bw=1



(b) Sub-humi - bw=2

In terms of the impacts by crop, in addition to the pattern in lagged rainfall differences in impacts could be partially explained by the differential in terms of technology adoption. We notice that, in the case of rice and wheat, where increases in irrigation and in the adoption of improved varieties from the 1960s to the 1990s were highest, the decrease in drought impacts was more pronounced in this period, compared to crops which are less often grown under irrigated conditions (sorghum, millet and maize). Since the beginning of the millennium, one plausible explanation as to why the drought coefficient for rice was less severely affected could lie in the fact that the decrease in lagged rainfall seems to have been less severe in sub-humid areas, where most of the rice-producing districts lie.

6.5 Conclusion

We argue that rainfall thresholds, used to define extreme rainfall events, should be intrinsically tied to tangible outcomes such as agricultural productivity. This is necessary in order to characterize the level of deviations from average rainfall levels beyond which additional deviations directly translate into significant impacts. Thus, we identified thresholds beyond which excess rainfall or rainfall deficiency begin to have negative impacts on cereal productivity. We then estimated the evolution of impacts over time across agro-ecological zones and crops.

Overall, we find five results of interest. First, for India as a whole, we find that rainfall deficiency starts having negative impacts on cereal productivity above the level used by the government for triggering its response to a drought event.

Second, there is very large heterogeneity in the levels of rainfall at which we identify these thresholds by agro-ecological zone. Broadly, on average, impacts become negative at smaller deviations from LTAR in humid and sub-humid areas, although impacts tend to be smaller in magnitude. In more arid areas, on the other hand, thresholds occur later (potentially due to adaptation), but impacts are more severe.

Third, we corroborate the prevailing consensus regarding the belief that crops such as millet, sorghum and maize tend to be more tolerant to rainfall deviations. However, we also find that at very large levels of rainfall deficiency they exhibit very large impacts compared to, for instance, wheat. We argue that this is likely to be due to the fact that such crops are almost entirely rainfed and as such, beyond a certain point their impacts are large, despite their tolerance for water stress.

Fourth, our method overwhelmingly finds thresholds below the LTAR, although this does not mean that extreme excess rainfall events are unimportant. Using an alternative method, which does not account for temperature, we find such events do indeed have negative impacts in semi-arid and sub-humid areas at very extreme levels of excess rainfall, generally beyond the 95th percentile, although impacts are typically smaller. A similar pattern applies for the crop analysis, where we also find significant negative impacts for sorghum and maize for excessive rainfall, although these occur earlier than for the agro-ecological estimates. However, since these events, for the most part, occur at very extreme values of the threshold regression, they are unlikely to be detected by the threshold approach.

Finally, we also analyse the evolution of impacts over time. Overall, we find that, while impacts decreased up until the late nineties, there seems to have been a reversal of the pattern since the beginning of the millennium. This result seems quite consistent across crops and agro-ecological zones, although crops such as rice have a smaller increase in impacts since the beginning of the millennium. The decrease in impacts until the late nineties is likely to be attributable to factors such as the increase in irrigation and adoption of improved varieties. The reversal could be explained by a noticeable change in weather patterns, whereby we notice that drought affected districts tend to have lower precipitation in the year preceding the drought and this pattern is particularly stark since the beginning of the millennium.

Our results highlight two aspects. First, they indicate that for a spatial unit the size of India, a unique threshold of rainfall deficiency and/or excess does not find any support in the data in terms of its impacts on agricultural productivity. These patterns of agro-ecological vulnerability together with changing patterns of precipitation and temperature at the regional scale are likely to be important for researchers examining the potential distributional impacts of climate change scenarios. Since we find that different crops have radically different tolerances to different thresholds, crop choice may have important implications in terms of adaptation to climate change. Our results suggest that the impacts of rainfall deficiency and excess rainfall are very asymmetric. With the exception of arid areas, our results support optimal ranges of rainfall close to the LTAR. They also suggest that while relatively small negative deviations may have large effects on productivity, only very large positive deviations in rainfall deviations have significant and large impacts on crop productivity. Second, they also indicate that while impacts of drought on yields have become smaller, their evolution has been highly non-linear.

This research has important implications with respect to climate change projections. We show that different spatial distribution of rainfall will have very different effects. Marginal impacts are largest for arid and semi-arid areas, when there are very low levels of rainfall. As such, if climate change drives rainfall patterns such that a large number of such events occur in arid areas, effects will be large. However, if effects are mainly felt in humid areas, overall, the effects are likely to be a lot smaller.

References

[1] Auffhammer, M., Ramanathan, V. & Vincent, J. R. (2012). Climate Change, the Monsoon, and Rice Yield in India. *Climatic Change*, 111(2):411-424.

[2] Birthal, P., Negi, D., Khan, M., & Agarwal, S. (2015). Is Indian Agriculture Becoming Resilient to Droughts? Evidence from Rice Production Systems. *Food Policy*, 56, 1-12.

[3] Burgess, R., Deschenes, O., Donaldson, D. & Greenstone, M. (2014). The Unequal Effects of Weather and Climate Change: Evidence from Mortality in India.

[4] Collier, M., & Webb, R. H. (2002). *Floods, droughts, and climate change*. University of Arizona Press.

[5] De, U. S., Dube, R. K., & Rao, G. P. (2005). Extreme weather events over India in the last 100 years. *Journal of the Indian Geophysical Union*, 9(3), 173-187.

[6] Deryng, D., Conway, D., Ramankutty, N., Price, J., & Warren, R. (2014). Global crop yield response to extreme heat stress under multiple climate change futures. *Environmental Research Letters*, 9(3), 034011.

[7] Diffenbaugh, N., Swain, D. & Touma, D. (2015). Anthropogenic Warming Has Increased Drought Risk in California. *Proceedings of the National Academy of Sciences of the United States of America*, 112(13), 3931-3936.

[8] Easterling, D. R., Meehl, G. A., Parmesan, C., Changnon, S. A., Karl, T. R., & Mearns, L. O. (2000). Climate extremes: observations, modeling, and impacts. *science*, 289(5487), 2068-2074.

[9] FAO (2012). *Crop Yield Response to Water*. Technical report, Food and Agriculture Organization of the United Nations.

[10] Gadgil, S. & Gadgil, S. (2006). The Indian Monsoon, GDP and Agriculture. *Economic and Political Weekly*, 41(47), 86-91.

[11] Gajbhiye, K. S., & Mandal, C. (2000). Agro-ecological zones, their soil resource and cropping systems. *Status of Farm Mechanization in India, cropping systems, status of farm mechanization in India*, 1-32.

[12] Ghosh, S., Das, D., Kao, S. C., & Ganguly, A. R. (2012). Lack of uniform trends but increasing spatial variability in observed Indian rainfall extremes. *Nature Climate Change*, 2(2), 86-91.

[13] Goswami, B. N., Venugopal, V., Sengupta, D., Madhusoodanan, M. S., & Xavier, P. K. (2006). Increasing trend of extreme rain events over India in a warming environment. *Science*, 314(5804), 1442-1445.

[14] Guiteras, R. (2009). The impact of climate change on Indian agriculture. *Manuscript, Department of Economics, University of Maryland, College Park, Maryland.*

[15] Hansen, B. E. (1999). Threshold effects in non-dynamic panels: Estimation, testing, and inference. *Journal of econometrics*, 93(2), 345-368.

[16] Hansen, B. E. (2000). Sample splitting and threshold estimation. *Econometrica*, 68(3), 575-603.

[17] Hatfield, J. L., Boote, K. J., Kimball, B. A., Ziska, L. H. & Izaurralde, R. C. (2011). Climate Impacts on Agriculture: Implications for Crop Production. *Agronomy Journal*, 103(2), 351-370.

[18] ICRISAT, (2012). District Level Database Documentation. Technical report, ICRISAT-ICAR-IRRI Collaborative Research Project.

[19] IPCC, (2012). Field, C. B. (Ed.). (2012). *Managing the risks of extreme events and disasters to advance climate change adaptation: special report of the intergovernmental panel on climate change*. Cambridge University Press.

[20] Indian Metereological Department, Government of India (undated). Frequently Asked Questions: FAQ. Available at: here

[21] Kumar, K. N., Rajeevan, M., Pai, D. S., Srivastava, A. K., & Preethi, B. (2013). On the

observed variability of monsoon droughts over India. Weather and Climate Extremes, 1, 42-50.

[22] Mishra, A. K. & Singh, V. P. (2010). A Review of Drought Concepts. *Journal of Hydrology*, 391(1), 202-216.

[23] Lobell, D. B., Sibley, A., & Ortiz-Monasterio, J. I. (2012). Extreme heat effects on wheat senescence in India. *Nature Climate Change*, 2(3), 186-189.

[24] Lobell, D. B., & Gourdji, S. M. (2012). The influence of climate change on global crop productivity. *Plant Physiology*, 160(4), 1686-1697.

[25] Pai, D. S., Sridhar, L., Rajeevan, M., Sreejith, O. P., Satbhai, N. S., & Mukhopadhyay, B. (2014). Development of a new high spatial resolution (0.25 0.25) long period (19012010) daily gridded rainfall data set over India and its comparison with existing data sets over the region. *Mausam*, 65(1), 1-18.

[26] Pai, D. S., Sridhar, L., Guhathakurta, P., & Hatwar, H. R. (2011). District-wide drought climatology of the southwest monsoon season over India based on standardized precipitation index (SPI). *Natural hazards*, 59(3), 1797-1813.

[27] Bhandari, H., Pandey, S., Sharan, R., Naik, D., Hirway, I., Taunk, S. K., & Sastri, A. S. R. A. S. (2007). Economic costs of drought and rice farmers drought-coping mechanisms in eastern India. *Economic Costs of Drought and Rice Farmers Coping Mechanisms: A Cross-Country Comparative Analysis, International Rice Research Institute, Los Baos, Philippines*, 43-111.

[28] Prasad, P. V. V., Staggenborg, S. A., & Ristic, Z. (2008). Impacts of drought and/or heat stress on physiological, developmental, growth, and yield processes of crop plants. Response of crops to limited water: *Understanding and modeling water stress effects on plant growth processes*, 301-355.

[29] Schlenker, W., & Roberts, M. J. (2009). Nonlinear temperature effects indicate severe damages to US crop yields under climate change. *Proceedings of the National Academy of sciences*, 106(37), 15594-15598.

[30] Shah, T., & Kishore, A. (2009). Will the impact of the 2009 drought be different from 2002?. Economic and Political Weekly, 11-14.

[31] Sharma, K. D. (2011). Rain-fed agriculture could meet the challenges of food security in India. *Current Science*, 100(11), 1615-1616.

[32] Singh, D., Tsiang, M., Rajaratnam, B., & Diffenbaugh, N. S. (2014). Observed changes in extreme wet and dry spells during the South Asian summer monsoon season. *Nature Climate Change*, 4(6), 456-461.

[33] Singh, A., Phadke, V. S., & Patwardhan, A. (2011). Impact of drought and flood on Indian food grain production. In *Challenges and Opportunities in Agrometeorology* (pp. 421-433). Springer Berlin Heidelberg.

[34] Srivastava, A. K., Rajeevan, M., & Kshirsagar, S. R. (2009). Development of a high resolution daily gridded temperature data set (19692005) for the Indian region. *Atmospheric Science Letters*, 10(4), 249-254.

[35] Turner, A. G., & Annamalai, H. (2012). Climate change and the South Asian summer monsoon. *Nature Climate Change*, 2(8), 587-595.

[36] Vicente-Serrano, S. M., Lopez-Moreno, J. I., Beguera, S., Lorenzo-Lacruz, J., Sanchez-Lorenzo, A., Garca-Ruiz, J. M., Azorin-Molina, C., Moran-Tejeda, E., Revuelto, J., Trigo, R. & Coelho, F. (2014). Evidence of increasing drought severity caused by temperature rise in southern Europe. *Environmental Research Letters*, 9(4), 044001.

[37] Wang, Q. (2015). Fixed-effect panel threshold model using Stata. *Stata Journal*, 15(1), 121-134.

[38] Wilhite, D. A. (2000). *Drought: A Global Assessment*. Routledge, London.

[39] Wilhite, D. A., & Glantz, M. H. (1985). Understanding: the drought phenomenon: the role of definitions. *Water international*, 10(3), 111-120.

[40] Yu, T., & Babcock, B. A. (2010). Are US corn and soybeans becoming more drought tolerant?. *American Journal of Agricultural Economics*, 92(5), 1310-1323.

Chapter 7

Conclusion - Summary of findings, implications and avenues for future research

Chapter 2 used a different methodology to compute the efficiency scores of farmers in Ethiopia. In doing so, it showed that relaxing the assumption of a single frontier leads to higher efficiency scores. As a result, this suggests that part of the high inefficiencies typically found in the literature may be a result of the methods used to compute them. This is not innocuous, since it implies that simply tackling agricultural inefficiencies is unlikely to be enough to deliver the desired increases in cereal production. Moreover, given the different elasticities estimated for different groups, using this method also allows recognizes the diverse needs of different farmers.

However, although the latent-class model allows for different groups of farmers to be compared to different frontiers, it is not devoid of problems. Chief among these, it introduces the issue of understanding what factors define a class of farmers. In the paper, the aim was to divide classes based on inputs to production. However, while I argue that this improves on conventional approaches which assume a unique frontier, the variables I propose for the latent-class allocation are not the only possible choice. Similarly, the estimation in Chapter 2, in line with the literature, imposes fixed classes over time. This implies that households cannot transition from one class to another during the sample period. In a long panel such as the one used in chapter 2, it could be argued that allowing households to switch classes is plausible.

As a result, there are at least two fruitful avenues for research that stem from Chapter 2. The first is to assess whether the result found in this paper is confined to this dataset, or whether it emerges in other settings. Second, further research is needed with regards to additional variables that would improve the class allocation of farmers.

Chapter 3 revisited the link between cereal diversity and production. In line with previous literature on the topic, we initially find a positive correlation between cereal diversity and cereal production. However, this relation becomes a lot weaker when households cultivating a low-productivity crop (teff) are removed from the sample. This suggests that, in our sample, yield differentials across cereals could partly explain the positive relationship found between production and diversity. This finding has far-reaching policy implications. A large effect of cereal diversity on cereal production would suggest that more diverse cereal production systems could potentially lead to large increases in food production. Our results, however, highlight that this may not be the case.

Nevertheless, our analysis suffers from two important limitations, which imply that we are not able to convincingly rule out whether a positive correlation exists or not. First, while we attempt to limit the concerns surrounding endogeneity (by using household fixed effects and village-year fixed effects), without a convincing instrument, we cannot rule out the issue of endogeneity. Secondly, our finding is plausible in our dataset since we focus on different cereals. However, some studies within the crop diversity literature have shown a similar effect to exist, even in the case of intra-cereal diversity (for different varieties of the same cereal). Unfortunately, we do not know enough about the context or the different types of a given cereal (typically wheat) in those samples to infer whether the channel we suggest plausibly explains those results as well. It is possible that difference in yields for different subspecies of a crop could explain that relationship.

Two potential paths emerge for future research. First, it would be a big step forward within the crop-diversity literature if a plausible instrument was found for crop diversity, since the issue of endogeneity would be resolved. Second, the literature has tended to have a strong focus on quantifying the effect of crop diversity. There has been less research on attempting to disentangle where the effect comes from (sampling, complementarity, facilitation or yield differential). Understanding the drivers of the impacts is important as different sources of the impact could have very different policy implications.

Chapter 4 estimates the impact of adopting SWC technologies on adult and child labour and finds that there is a consistently large and significant impact of adoption on adult labour. A similar impact is found for child labour, although it is only statistically significant in some specifications?.

There are three implications that stem from these results. First, at the macro-level, these estimates of labour impacts can be useful for CGE modellers who focus on the general equilibrium effects of the adoption of SWC technologies. After all, agriculture still accounts for approximately 75% of the labour force in Ethiopia. As a result, widespread promotion of labour-intensive technologies can have non-negligible general equilibrium effects. Secondly, the suggested increases in child labour is also something that policy-makers can anticipate and take into consideration when promoting such technologies. Governments could, for example, promote these technologies while, at the same time, increasing the incentives for alternative uses of childrens time (e.g. Conditional Cash Transfers where school attendance is a condition). Third, these results should serve as a warning for practitioners estimating impacts on production using cross-sectional data and methods that assume no change in inputs resulting from adoption.

There are two main implications for future research. First, this paper highlights the need to move towards panel data and to methods that allow researchers to control for these changes in inputs (e.g. the Semi-Parametric DID proposed by Abadie) when estimating impacts on production. Second, it highlights the need for analysing impacts of technology adoption on outcomes other than production and productivity.

Chapter 5 first reviews the suitability of the drought index proposed by Babcock and Yu (2010). We argue that the index these authors propose is too narrow in its coverage of dry events. Babcock and Yu (2010) define droughts as events with below-average precipitation and above-average temperature (“hot droughts). We extend and modify their index to account for events with below-average precipitation and below-average temperature (“cold droughts). Our opinion is that these events should not be dismissed a priori and their importance should be tested. We find that cold droughts have a large and statistically significant impact on production and that their exclusion leads to an underestimation of drought impacts.

One implication that emerges from this paper is that it is important to use a drought index that encompasses all potential dry events. In principle, indices that ignore cold droughts should be avoided without prior testing. However, if for some reason only "hot droughts are of interest in a given setting, a dummy variable should be included to derive more accurate marginal effects. We also argue that indices that use arbitrary precipitation thresholds are likely to be inaccurate (this is further explored in chapter 6) as they do not incorporate temperature and may not capture negative impacts prior to that threshold.

Chapter 6 focuses on the impact of extreme rainfall events attempts to answer two questions. First, it identifies thresholds of rainfall parametrically. Then, we complement the threshold results by estimating impacts non-parametrically, using rainfall dummies for different percentiles. Second, focusing on events of rainfall deficiency, it looks at how the impacts of drought have evolved over time in India.

A number of results emerge from this analysis. First, we identify empirically determined ranges of rainfall for which the marginal impacts of drought are different. We notice that these thresholds differ substantially by agro-ecological zone and crop. Typically, we find that, in more arid areas and for more drought-resistant crops (millet, sorghum and maize), thresholds occur at lower levels of rainfall, but that beyond these thresholds the impacts of drought are very large. This means that impacts are small at low negative deviations from normal rainfall. However, they become very large beyond certain thresholds of rainfall. Second, using the non-parametric analysis, we find that there are also cases where excessive positive rainfall also has a negative impact on productivity, though typically the impacts are smaller and occur only at very extreme levels of rainfall (generally beyond the 90th percentile of rainfall). Finally, with respect to the evolution of drought impacts over time, we find a decrease in drought impacts until the late 1990s and a reversal of this trend thereafter. We postulate that this may be due to a change in the patterns of rainfall in drought-affected districts since the late 1990s.

This paper also suffers from a number of limitations. First, the threshold model is useful to detect where the marginal impacts of the RTI change. However, this may not necessarily coincide with the levels of rainfall at which impacts become negative. Second, the threshold model used in this paper is unlikely to identify thresholds that are very close to the end of the distribution of rainfall as the estimated coefficients are likely to be noisy. As a result, we use a complementary method that models the impacts of precipitation using a series of dummy variables. The latter, however, does not incorporate temperature in the estimation. A third weakness of the threshold model is that it is not able to estimate the evolution of the threshold over time though it is not clear if this would be desirable. It is plausible that, as a result of technological change, negative thresholds now occur later. However, we would be unable to pinpoint the source of the change in the threshold, whether it is a result of technological change or whether it stems from a change in the relative importance of rainfall and temperature over time. In the time results, rolling regressions can be criticized on the grounds that they capture a mixture of the true coefficient and noise.

There are four policy implications that emerge from this paper. First, the paper highlights the dangers of using arbitrarily defined thresholds. We often find significant negative impacts above those thresholds. Second, the paper highlights the danger of treating India as a homogeneous unit. India is a very diverse country and results for the full sample cannot necessarily be extrapolated for every area and/or every crop. Third, and most importantly, the threshold and non-parametric estimates could allow policy-makers to dose their responses according to the severity of the rainfall deficiency rather than allowing an arbitrary threshold to trigger a binary response system. Fourth, our paper shows that India is not drought-proof and, although there is some evidence that impacts have decreased over time, this positive trend has been reversed of late.

These results point to at least three areas for future research. First, deriving a methodology that would allow researchers to analyse the evolution of thresholds over time, while accounting for composition effects of the drought index, would be interesting. Second, it would also be interesting to understand whether the reversal in the trends seen in the rolling regressions has continued since 2009 or whether it has reversed once again. The latter is more plausible as it is likely that the increase in impacts after the millennium is driven by the years 1999-2006. Finally, it would be interesting to better investigate the causes leading to the increases in drought impacts since the beginning of the millennium, as we are not able to prove beyond dispute that the change in rainfall patterns for drought-affected districts is the cause of this.