Technische Universität München
TUM School of Life Sciences

Agricultural Production and Environmental Change: An Economic Investigation

Christian Wendelin Stetter

Vollständiger Abdruck der von der TUM School of Life Sciences der Technischen Universität München zur Erlangung eines

Doktors der Agrarwissenschaften (Dr.agr.)

genehmigten Dissertation.

Vorsitz: Prof. Dr. Anja Rammig

Prüfer*innen der Dissertation:

1. Prof. Dr. Johannes Sauer
2. Prof. Dr. Thomas Knoke
3. Prof. Dr. Jochen Kantelhardt

Die Dissertation wurde am 05.07.2022 bei der Technischen Universität München eingereicht und durch die TUM School of Life Sciences am 15.11.2022 angenommen.

Acknowledgements

First of all, I would like to thank the examining committee of this thesis for having dedicated their time to reading and reviewing this work.

This thesis would not have been possible without the combined support of many people. I would like to thank my supervisor, Prof. Dr. Johannes Sauer for trusting my research abilities, giving me the liberty to develop and implement own ideas, and for interesting discussions and challenging feedback. This helped me a lot to grow as a researcher.

Further, a special thanks goes to my former and current colleagues at the chair group of Production and Resource Economics at the TU Munich. I am also grateful to my co-authors Dr. Stefan Wimmer and Philipp Mennig for great collaborations and discussions as well as to my mentor Dr. Fabian Frick for his continuous feedback and the open exchange. I am very grateful to Katherine and Tom O'Bryan for having taken the time to proofread parts of this thesis.

I am also indebted to my family, especially to my parents for their continuous and invaluable support. Finally, a huge thank you goes to my wife, Amira, for supporting me during this doctoral project and all the years before, for bearing with me during a two-year home office period, and for always having a sympathetic ear for me.

Summary

The benefits of agriculture have been enormous and it has experienced a large increase in productivity over the last decades. It is the source of the world's food supply. At the same time, farming activities have been linked to a series of environmental problems such as climate change, biodiversity loss, and soil degradation. In this context, climate change has been playing an important role in recent years. It is expected to drastically modify the natural conditions under which farmers produce. In the face of a changing environment and an increasing demand for food, energy and renewable resources, agricultural production is facing the twofold challenge of increasing production without degrading the environment (and the climate). To reconcile the two objectives of an increased productivity and environmental friendliness, farms have to mitigate their adverse impacts on the environment and the climate. At the same time, they must adapt to prevailing environmental and climatic changes.

This thesis aims to provide novel insights into the relationship between agricultural production and environmental change at the micro-level by combining microeconomic theory with state-of-the-art econometric techniques. It is comprised of four empirical studies and focuses geographically on Bavaria, a federal state of Germany and one of the core agricultural regions in Europe.

Study 1 develops a parametric eco-efficiency concept capable of jointly evaluating the ecological and economic performance of (farm) businesses over time and empirically applies this concept to greenhouse gas emissions on farms. It distinguishes between persistent (long-term) and time-varying (short-term) efficiency. The study finds rather small levels of time-varying emission-inefficiency and high levels of persistent inefficiency across different farm types. Overall, the sample farms were mostly not emission-efficient. The parametric stochastic frontier approach allowed to capture eco-performance dynamics over time based on a generalized Malmquist productivity index. Emission performance slightly improved between 2005 and 2014 for most farm types.

Study 2 measures the environmental and economic performance of extensive and intensive dairy farms with respect to greenhouse gas emissions by combining the concept of eco-efficiency with latent-class stochastic frontier analysis and the estimation of a stochastic meta-frontier. It finds that intensive farms were on average more efficient at minimizing greenhouse gas emissions at given economic return levels than their extensive counterparts. This study shows that technology differences matter with respect to ecological-economic performance. Overall, there is large potential for climate change mitigation without risking economic viability for both intensive and extensive farms.

Study 3 assesses the heterogeneous impacts of agri-environmental schemes on the environmental performance of farming. The study combines economic theory with a novel machine learning method to identify the environmental effectiveness of agri-environmental schemes at the farm level. Results from the empirical application suggest the existence of heterogeneous, but limited effects of agri-environment measures across several environmental dimensions such as climate change mitigation, clean water and soil health. By making use of model-agnostic interpretation methods, the importance of production heterogeneity and context-specificity in agricultural policy evaluation is highlighted and important insights into the improved targeting of agri-environmental schemes using contextual knowledge are demonstrated.

Study 4 evaluates whether, and under what conditions, farmers are likely to adopt agroforestry and wood-based land-use systems in response to regional weather extremes. The cultivation of agroforestry systems is regarded as an effective strategy to synergistically mitigate and adapt to the impacts of climate (and environmental) change in the face of increased occurrences of regional extreme weather events. The study combines a discrete choice experiment with geospatial weather information. Assuming adaptive weather expectations, land users' dynamic responses to extreme weather years are simulated in terms of adoption probabilities. The results from the simulation of a 2018-like extreme weather year suggest that agroforestry might have a very high probability of being adopted in the medium to long-term under different scenarios, thus enhancing farmers' resilience to climate change.

The findings of this thesis contribute to the growing literature on the production-environment nexus in agriculture. The findings of Study I and Study II demonstrate that economic success and environmental (and climate) protection do not necessarily have to be mutually exclusive. Furthermore, the methodological steps (including causal machine learning and model-agnostic interpretation) and findings of Study III help to overcome limitations in the assessment of agri-environmental schemes and open new possibilities for agricultural impact assessments. The findings of Study IV add to a better understanding of climate (and environmental) change adaptation and mitigation in the face of extreme weather events.

Ultimately, a number of cross-cutting themes are discussed, including the role of farm system dynamics, production heterogeneity and context specificity, as well as data availability and requirements. The thesis closes with an in-depth discussion of policy implications with a special emphasis on agri-environmental schemes and delineates pathways for future research.

Zusammenfassung

Die Leistungstärke landwirtschaftlicher Betriebe zur Herstellung pflanzlicher und tierscher Produkte ist enorm und erlebte starke Produktivitätsteigerungen in den vergangenen Jahrzehnten. Landwirtschaftliche Aktivitäten werden allerdings mit einer Reihe von Umweltproblemen wie Klimawandel, Verlust der biologischen Vielfalt und Bodendegradation in Verbindung gebracht. Darüber hinaus, ist die Landwirtschaft stark von der Umwelt selbst abhängig. In diesem Zusammenhang spielt der Klimawandel seit einigen Jahren eine außerordentliche Rolle. Es wird erwartet, dass der Klimawandel die natürlichen Bedingungen, unter denen Landwirte produzieren, drastisch verändern wird. Angesichts einer sich verändernden Umwelt und einer steigenden Nachfrage nach Nahrungsmitteln, Energie und nachwachsenden Rohstoffen steht die landwirtschaftliche Produktion vor der Herausforderung, die Produktion zu steigern, ohne dabei jedoch die Umwelt (und das Klima) zu beschädigen. Um die beiden Ziele, Produktivitätssteigerung und Umweltfreundlichkeit, in Einklang zu bringen, müssen landwirtschaftliche Betriebe ihre negativen Auswirkungen auf Umwelt und Klima mindern. Gleichzeitig müssen sie sich an vorherrschenden Umwelt- und Klimaveränderungen anpassen.

Die vorliegende Dissertation zielt darauf ab, neue Einblicke in die Beziehung zwischen landwirtschaftlicher Produktion und der Veränderung der Umwelt auf Betriebsebene zu liefern, indem sie mikroökonomische Theorie mit modernsten ökonometrischen Techniken kombiniert. Die Thesis besteht aus vier empirischen Studien, die sich räumlich auf den Freistaat Bayern konzentrieren, ein Bundesland im Südosten Deutschlands, welches eine der landwirtschaftlichen Kernregionen in Europa darstellt.

In Studie I wird ein parametrisches Ökoeffizienz-Konzept entwickelt, welches gleichzeitig die ökologische und ökonomische Leistung von (landwirtschaftlichen) Betrieben bewertet. Das entwickelte Konzept wird dann empirisch auf Treibhausgasemissionen landwirtschaftlicher Betriebe angewendet. Dabei wird zwischen persistenter (langfristiger) und zeitvariabler (kurzfristiger) Effizienz unterschieden. Die Betrie-

be in der bayerischen Fallstudie zeigten eine geringes Niveau an zeitvariabler Ineffizienz und ein hohes Maß an persistenter Ineffizienz für verschiedene Betriebstypen. Insgesamt waren die untersuchten Betriebe meist nicht sehr emissions-effizient. Der verwendete Stochastik Frontier-Ansatz ermöglicht es, Öko-Performance im Verlauf der Zeit auf der Grundlage eines Malmquist-Produktivitätsindex zu erfassen. Insgesamt hat sich die Öko-Performance bezogen auf den Klimawandel zwischen 2005 und 2014 leicht verbessert.

In Studie II wird die ökologische und wirtschaftliche Leistung extensiver und intensiver Milchviehbetriebe in Bezug auf Treibhausgasemissionen gemessen, indem das Konzept der Ökoeffizienz mit einer Stochastik Frontier-Analyse und der Latenten Klassenanalyse sowie der Schätzung einer stochastischen Metafrontier kombiniert wird. Die Ergebnisse zeigen, dass intensiv wirtschaftende landwirtschaftliche Betriebe im Durchschnitt effizienter Treibhausgasemissionen (bei gegebenen wirtschaftlichen Erlösen) als extensive Betriebe minimiert haben. Diese Studie zeigt, dass Technologieunterschiede im Hinblick auf die ökologisch-ökonomische Leistungfähigkeit von Milchviehbetrieben eine wichtige Rolle spielen. Insgesamt besteht ein großes Potenzial zum Klimaschutz, ohne dabei die Wirtschaftlichkeit zu gefährden. Das gilt sowohl für intensiv als auch extensiv wirtschaftende landwirtschaftliche Betriebe.

In Studie III werden die heterogenen Auswirkungen von Agrarumweltprogrammen auf die Umweltleistungen landwirtschaftlicher Betriebe untersucht. Diese Studie kombiniert Grundlagen ökonomischer Theorie mit Methoden maschinellen Lernens, um die Umweltwirksamkeit von Agrarumweltprogrammen auf der einzelbetrieblichen Ebene zu analysieren. Ergebnisse aus der empirischen Anwendung deuten auf heterogene, jedoch begrenzte Effekte von Agrarumweltmaßnahmen über mehrere Umweltdimensionen wie Klimaschutz, sauberes Wasser und Bodengesundheit hinweg hin. Durch die Verwendung modellagnostischer Interpreationsmethoden wird die Bedeutung von Produktionsheterogenität und Kontextspezifität in der agrarpolitischen Bewertung solcher Programme hervorgehoben und wichtige Einblicke in ein verbessertes Targeting analysiert.

In Studie IV wird bewertet, ob und unter welchen Bedingungen Landwirte als Reaktion auf regionale Wetterextreme Agroforstwirtschaft und holzbasierte Landnutzungssysteme einführen. Der Anbau von Agroforstsystemen gilt als wirksame Strategie zur synergetischen Minderung und Anpassung an die Auswirkungen des

Klimawandels, unter anderem im Hinblick auf das vermehrte Auftreten regionaler Extremwetterereignisse. Die Studie kombiniert ein Discrete-Choice-Experiment mit Geo-Wetterdaten. Unter der Annahme adaptiver Wettererwartungen werden die Reaktionen von Landnutzern auf Extremwetterjahre im Hinblick auf Anbauwahrscheinlichkeiten simuliert. Die Ergebnisse aus der Simulation eines Extremwetterjahres zeigten, dass Agroforstsysteme über verschiedene Szenarien hinweg mittel- bis langfristig mit sehr hoher Wahrscheinlichkeit eingeführt werden könnten, wodurch die Resilienz von Landwirten gegenüber dem Klimawandel erhöht wird.

Die Ergebnisse dieser Dissertation ergänzen die wachsende Fachliteratur zum Verhältnis von Produktion und Umwelt in der Landwirtschaft. Die Ergebnisse von Studie I und Studie II zeigen, dass sich wirtschaftlicher Erfolg und Umwelt- und Klimaschutz nicht zwangsläufig gegenseitig ausschließen müssen. Methodische Erkenntnisse (einschließlich kausalem maschinellem Lernen und modellagnostischer Interpretation) und die Resultate aus Studie III helfen dabei, methodische Limitationen bei der Bewertung von Agrarumweltprogrammen zu überwinden und eröffnen neue Möglichkeiten für kausale Folgenabschätzungen in der Landwirtschaft. Die Ergebnisse von Studie IV tragen zu einem verbesserten Verständnis der Anpassung an den Klima- und Umweltwandel und dessen Minderung bei.

Letztendlich werden eine Reihe von Querschnittsthemen, darunter die dynamische Natur landwirtschaftlicher Systeme, Produktionsheterogenität und Kontextspezifität sowie Datenverfügbarkeit und -anforderungen diskutiert. Die Dissertation schließt mit einer Diskussion über politische Implikationen der Ergebnisse unter besonder Berücksichtigung von Agrarumweltprogrammen und skizziert zukünftigen Forschungsbedarf im Kontext dieser Arbeit.

Table of contents

List of figures

List of tables

Abbreviations

AC	alley cropping
AD	activity data
AES	agri-environment schemes
AIC	Akaike Information Criterion
ATE	average treatment effect
ATT	average treatment effect on the treated
AWU	average work unit
BACON	blocked adaptive computationally efficient outlier nominators
BIC	Bayesian Information Criterion
CaCO	calcium carbonate
CAP	Common Agricultural Policy of the EU
CART	classification and regression trees
CATE	conditional average treatment effect
$\mathbf{CH_4}$	methane
$\mathbf{CO_2}$	carbon dioxide
$\mathbf{CO_{2eq}}$	CO_2-equivalents
CRS	constant returns to scale
DAG	directed acyclic graph
DCE	discrete choice experiment
DEA	data envelopment analysis
DiD	difference-in-difference
DMU	decision-making unit
EE	emission efficiency
EEC	emission-efficiency change
EEF	eco-efficiency frontier
EF	emission factor
EP	eco-performance
EPC	eco-performance change with respect to emissions

ESC	emission scale change
ETC	ecological-technical change with respect to emissions
EU	European Union
FADN	Farm Accountancy Data Network
FAO	Food and Agriculture Organization of the United Nations
FSDN	Farm Sustainability Data Network
GHG	greenhouse gas
GRF	generalized random forest
ha	hectare
HTE	heterogeneous treatment effect
IPCC	Intergovernmental Panel on Climate Change
IR	iso-revenue line
KULAP	Bayerisches Kulturlandschaftsprogramm
LASSO	least absolute shrinkage and selection operator
LCA	life cycle analysis
LfL	Bavarian State Research Center for Agriculture
LU	livestock unit
ML	machine learning
MTE	metaforntier efficiency
N	nitrogen
$\mathbf{N_2O}$	nitrous oxide
NUTS	Nomenclature of Territorial Units for Statistics
OECD	Organisation for Economic Co-operation and Development
OLS	ordinary least squares
OVB	ommited variable bias
PC	principal component
PEE	persistent emission efficiency
PES	payments for ecosystem services
PGT	pressure-generating technology
PPC	production possibility curve
PPF	production possibility frontier
RDP	rural development program
REE	time-varying emission efficiency
RF	random forest
RPL	random parameter logit

RTES	returns to emission scale
SFA	stochastic frontier analysis
SRC	short rotation coppice
TE	technical efficiency
TFP	total factor productivity
TGR	technology gap ratio
UNFCCC	United Nations Framework Convention on Climate Change
VNP	Vetragsnaturschutzprogramm
VRS	variable returns to scale
WTA	willingness to adopt

Part I

Introduction and Methods

1 Introduction

1.1 Motivation

The benefits of agriculture have been enormous. Agricultural production activities are the source of the world's food supply. Worldwide, gross crop and livestock production has almost quadrupled since 1960 (FAOSTAT, 2022c). At the same time, the yields of the three most important staple crops (maize, rice, and wheat) have continuously increased (Figure 1.1a). This is also true for agriculture's total factor productivity (TFP), which has gone up by 81% since 1961 (Figure 1.1b). Furthermore, this development is also reflected by the amount of people that are fed per hectare of arable land, which increased from 2.4 people per hectare in 1961 to 5.6 people per hectare in 2019 (Figure 1.1c).

These developments over the past 60 years have guaranteed that there is sufficient food produced for nearly eight billion people (FAOSTAT, 2022c). Modern agriculture helped to reduce hunger, to improve nutrition and to mitigate extreme poverty (Tilman et al., 2002). Alongside these trends, profound changes in the environment could globally be observed (Tilman & Lehman, 2001). These changes have affected, among other things, the climate, biodiversity, soil, nutrient and carbon cycles, as well as ecosystem and land composition (Pyhälä et al., 2016). As earth is a dynamic system, environmental change has always been a fundamental part of its functioning (Vitousek, 1992). However, more recently changes in the environment could clearly be associated to human activities (Campbell et al., 2017; Tilman et al., 2001; Tilman & Lehman, 2001). This is particularly true for agriculture, which is inherently connected to environmental change. For instance, agricultural productivity gains over the past 60 years were mainly driven by a rise in inputs such as fertilizers, water and pesticides (see also Figure 1.2a), novel crop varieties, and other technological advances (Tilman et al., 2002). This development has coincided with a drastic increase in greenhouse gas (GHG) emissions (Figure 1.2b) and a decrease in farmland birds (Figure 1.2c).

A vast body of literature has found negative impacts of agricultural production

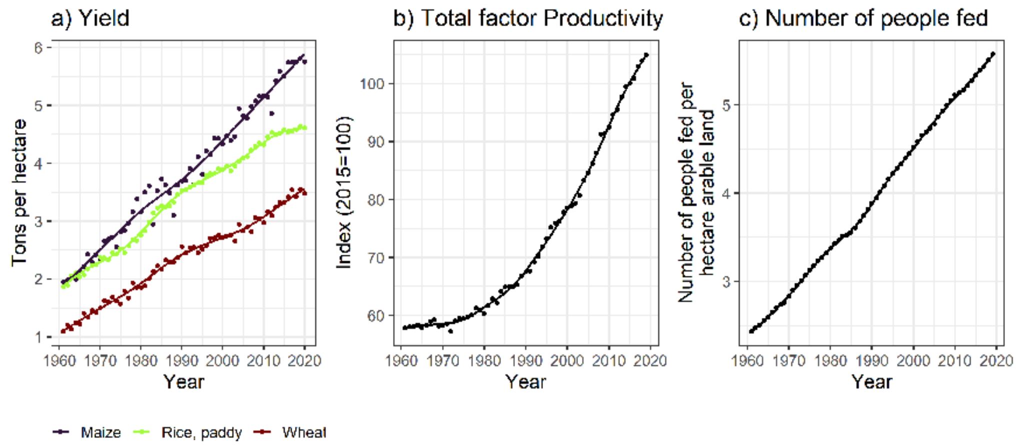


Figure 1.1: Productivity indicators on global agriculture (1960–2020). Sources: FAOSTAT (2022a,c,d) and USDA (2022).

activities on the environment (Pyhälä et al., 2016). For instance, current farming practices have been found to degrade soil due to erosion, salinization, acidification, contamination, or compaction (Kopittke et al., 2019; Tsiafouli et al., 2015). Adverse effects of farming have also been found on several biodiversity domains. Especially the agriculture-induced loss of insects reported in recent studies (Ewald et al., 2015; Gossner et al., 2016; Ramos et al., 2018; Seibold et al., 2019) has spurred an intense public debate. Further environmental problems associated with agricultural production include nitrate in groundwater (Scanlon et al., 2007), ammonia emissions (European Environment Agency, 2019b) and excessive pesticide use (European Environment Agency, 2018; Mahmood et al., 2016).

At the same time as agricultural production affects environmental conditions, it is also heavily dependent on the environment itself which provides important natural resources such as soils, water, genetic material, as well as ecosystem services such as soil fertility, pollination and pest control, that are crucial for production (Lichtenberg, 2002). Simultaneously, environmental degradation has already been shown to negatively influence farming activities (Zhang et al., 2007).

At the nexus of agricultural production and environmental change, climate change has arguably been playing an outstanding role in the field of agricultural economics in recent years. Climate change is expected to drastically modify the natural conditions under which farmers produce, and has a direct effect on agricultural production (Njuki et al., 2018; OECD, 2019). A growing body of evidence suggests that this effect is most likely detrimental. In this context, multiple studies have found statistically significant relationships between climate change and

TFP (Burke & Emerick, 2015; Chambers & Pieralli, 2020; Chambers et al., 2020; Mendelsohn et al., 1994; Njuki et al., 2020, 2018; Ortiz-Bobea et al., 2021, 2018; Schlenker et al., 2005), profits (e.g. Deschênes & Greenstone, 2007, 2012), labor (Burke et al., 2015), land use (Cui, 2020; Ramsey et al., 2021) and crop yields (e.g. Keane & Neal, 2020; Schlenker & Roberts, 2009; Vogel et al., 2019; Webber et al., 2020).

Given the current levels of world population growth in liaison with global trends regarding meat and dairy consumption, it is predicted that global food production will have to grow by up to another 35–56% by 2050 compared to 2010 (van Dijk et al., 2021). Furthermore, agriculture is expected to play a key role in the transformation of the world's economy toward a biobased future (Loiseau et al., 2016; Schmidt et al., 2012). Developing a bioeconomy will require an immense amount of fiber, energy, and renewable resources (see e.g. Biermann et al., 2011; Muffler & Ulber, 2008; Valentine et al., 2012). However, continuing farming as it has been done before will likely cause the earth system to exceed various planetary boundaries, which might have irreversible consequences for humankind's basis of life (Campbell et al., 2017; Rockström et al., 2009).

Hence, given agriculture's inherent multi-functionality and embedding in the environment, it is facing the two-fold challenge of increasing production without degrading the environment. Addressing this challenge heavily depends on what decisions farmers' make and what actions they take locally (Pyhälä et al., 2016). Thus, although this challenge constitutes a global endeavor, it is important to understand the decision-making processes and environmental change responses at the individual farm-level in that farmers are the agents or decision-making units in this context (Malek et al., 2019).

To reconcile the two objectives of an increased productivity and environmental friendliness, farms have to mitigate their adverse impacts on the environment and the climate (Rosenzweig & Tubiello, 2007). Simultaneously, they have to adapt to prevailing environmental and climatic changes (Zilberman et al., 2012). The scientific literature has suggested multiple instruments for both mitigation and adaptation. For instance, regarding mitigation, the Intergovernmental Panel on Climate Change (IPCC, 2014) listed restoration of organic soil as well as an adjustment of cropland and grassland management as effective measures. As for adaptation, suggestions from the literature range from conservation tillage (e.g.

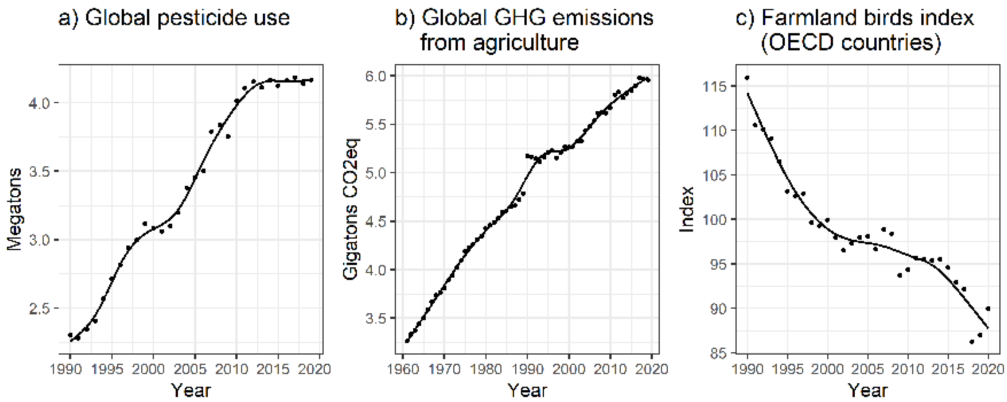


Figure 1.2: Environmental indicators associated with agricultural production (1960–2020). Sources: FAOSTAT (2022b,d); OECD (2022).

Iglesias et al., 2012) to irrigation (e.g. Klein et al., 2014) to insurance (e.g. Di Falco et al., 2014) and diversification (e.g. Olesen et al., 2011).

What is more, mitigation and adaptation in agriculture have been found to have great potential for synergies (Rosenzweig & Tubiello, 2007; Smith & Olesen, 2010). One potential measure in this context is the cultivation of agroforestry systems, which integrate woody perennials with agricultural crops and/or livestock on a piece of land (Verchot et al., 2007). They provide a large set of important ecosystem services, they are robust to climate change and extreme weather, and they can store atmospheric carbon (Brown et al., 2018).

Another important aspect worth mentioning in the context of agricultural production and environmental change is the role of agri-environmental policy instruments (Hasler et al., 2022). Legislation has been shown to contribute to reducing environmental problems/degradation caused by agricultural production (e.g. Arata & Sckokai, 2016; Bertoni et al., 2020; Chabé-Ferret & Subervie, 2013). However, an aspect that has largely been neglected so far, is the importance of considering the individual farming context to gain a better understanding of the heterogeneous impacts and effectiveness of agri-environmental policy measures (Dessart et al., 2019). One reason for that is the fact that the traditional econometric toolset available to agricultural and applied economists is limited in its ability to reflect the complexities and nonlinearities of the relationship between agricultural production and environmental change. However, recent advances in the causal machine learning literature have opened new possibilities to take such complexities into account (Storm et al., 2020).

1.2 Aims and scope of this thesis

This thesis aims to provide novel insights into the relationship between agricultural production and environmental change at the micro-level by combining microeconomic theory with state-of-the-art econometric techniques. Special consideration is given to three specific issues in this context. These are the trade-offs between farm-level economic returns and greenhouse gas emissions, the role of AES in promoting environmentally-friendly farming practices, and the potential of agroforestry adoption to tackle climate change through adaptation and mitigation. These issues were specifically investigated for farms in the case study region of Bavaria, a federal state in southeast Germany. A primary reason for choosing this region lies in its large diversity both in terms of its farm structures and in terms of its farming conditions, which may allow conclusions for various other regions beyond Bavaria as well.

The empirical part of this thesis is subdivided into four studies:

I. Greenhouse gas emissions and eco-performance at farm level: a parametric approach.

II. Are intensive farms more emission-efficient? Evidence from German dairy farms.

III. Using machine learning to identify heterogeneous impacts of agri-environment schemes in the EU: A case study.

IV. Tackling climate change: agroforestry adoption in the face of regional weather extremes.

with the following specific aims:

I. to develop a parametric eco-efficiency concept capable of jointly evaluating the ecological and economic performance of (farm) businesses over time and to empirically apply this concept to greenhouse gas emissions on Bavarian farms,

II. to measure the environmental and economic performance of extensive and intensive dairy farms with respect to greenhouse gas emissions by combining the concept of eco-efficiency with latent-class stochastic frontier analysis (SFA),

III. to assess the impact heterogeneity of agri-environmental schemes on the environmental performance of farms based on a novel causal machine-learning approach and to use contextual knowledge to improve scheme targeting,

IV. to evaluate whether, and under what conditions, farmers are likely to adopt agroforestry and wood-based land-use systems in response to regional weather extremes.

Farms can only sustainably produce agricultural goods and services when their businesses are economically viable. The relationship between economic viability of farms and its environmental impact can be looked at from an efficiency point of view, i.e. to ask the question by how much can GHG emissions be reduced without reducing the economic outcome of a farm, or conversely, by how much could economic returns be increased without increasing GHG emissions. In this context, the concept of ecological-economic efficiency, or "eco-efficiency" has gained recognition as a way of evaluating the balance between economic performance and the environmental damage induced by economic activity (Kuosmanen & Kortelainen, 2005). Yet, evidence on the trade-off between greenhouse gas emissions and economic performance is sparse. Chapters 3 and 4 encompass studies that assess this trade-off. Study I conceptually extends the eco-efficiency approach to a general stochastic frontier analysis (SFA) setting and develops an eco-performance indicator to better evaluate farms' performance with respect to GHG emissions over time. Study II relates this measure to separate production technologies and shows differences between intensive and extensive farm groups.

To reduce the environmental impact of farming activities, legislators introduced voluntary AES, in which farmers commit themselves to adopt environmentally-friendly farming techniques to mitigate adverse environmental effects and foster desirable ecosystem services. There is extensive literature on the effectiveness of such schemes. A major limitation of most studies is that they mostly look at average effects of such schemes, thereby neglecting the fact that AES are likely to affect different farms differently. Study III presented in chapter 5 uses recent methodological innovations from the causal machine learning literature to overcome this limitation and provide farm-level evidence on the environmental effectiveness of AES.

One major channel through which farmers can actively tackle environmental and climate change impacts is land use (Pielke, 2005). A promising pathway in this direction is the adoption of agroforestry and wood-based land-use systems, which are recognized to play a key role in synergistically approaching adaptation and mitigation (Verchot et al., 2007). Study IV presented in chapter 6 evaluates farmers probability of adopting such systems in the face of an increased number of extreme weather events.

The first three studies of this thesis have been published in international peer-reviewed scientific journals, and the fourth study is currently under review in such a journal. An overview of the studies can be found in Table 1.1.

1.3 Literature overview

This section provides a brief literature overview of applied research on the three dominant topics of this thesis: agricultural production and eco-efficiency, the effectiveness of AES, and agroforestry adoption in response to climate change.

1.3.1 Agricultural production from an eco-efficiency perspective

Over recent decades, the concept of ecological-economic efficiency, or "eco-efficiency" has gained recognition as a way of evaluating the balance between economic performance and the environmental damage induced by economic activity. The first studies concerned with eco-efficiency computed partial measures such as "economic output per unit of waste" (see Tyteca, 1996). Starting with the article by Kuosmanen & Kortelainen (2005), who operationalized the eco-efficiency concept to be calculated by means of data envelopment analysis (DEA), numerous studies in several fields followed, both at the micro- and at the macro-level.

Studies focusing on the assessment of eco-efficiency in agriculture have analyzed various farm types, ranging from crop farms (Bonfiglio et al., 2017; Eder et al., 2021; Gadanakis et al., 2015; Song & Chen, 2019; Yang et al., 2021) and dairy farms (Cortés et al., 2021; Iribarren et al., 2011; Martinsson & Hansson, 2021; Mu et al., 2018; Orea & Wall, 2017; Pérez Urdiales et al., 2016; Soliman & Djanibekov, 2021; Soteriades et al., 2016, 2020) to horticultural farms (Angulo-Meza et al., 2019; Godoy-Durán et al., 2017), olive (Beltrán-Esteve et al., 2014; Picazo-Tadeo et al., 2012) and wine producers (Grassauer et al., 2021). The majority of studies have a spatial focus on Europe including, e.g., Spain (e.g. Picazo-Tadeo et al., 2011), the UK (e.g. Gadanakis et al., 2015), Italy (e.g. Bonfiglio et al., 2017), Sweden (Martinsson & Hansson, 2021), Austria (e.g. Eder et al., 2021), and Poland (e.g. Pishgar-Komleh et al., 2020). Further analyses have been conducted, e.g., in New

Table 1.1: Overview of the empirical studies presented in Part II.

Study/ chapter	Research problem	Data	Methods	Core findings
Empirical study 1/ chapter 3	Joint evaluation of the ecological and economic performance of farms over time	Bavarian farms by farm types: dairy, pig, mixed, crop farms; financial accountancy data; 2005-2014	Eco-efficiency; SFA with persistent and time-varying eco-inefficiency; Malmquist productivity index decomposition	Little time-varying emission-inefficiency and high levels of persistent inefficiency; farms were mostly not very emission-efficient; eco-performance slightly improved, farm-type differences
Empirical study 2/ chapter 4	Comparison of the joint ecological and economic performance of extensive and intensive dairy farms with respect to greenhouse gas emissions	Bavarian specialized dairy farms, financial accountancy data; 2005-2014	Combination of eco-efficiency, latent-class SFA and stochastic meta-frontier analysis	Intensive farms were on average more efficient at minimizing GHG emissions at given economic return levels than extensive ones, large potential for climate change mitigation without risking economic viability
Empirical study 3/ chapter 5	Assessment of the impact heterogeneity of AES on the environment and improvment of scheme targeting	Bavarian farms (all types), financial accountancy data; 2014	Neyman–Rubin causal model (Neyman, 1923; Rubin, 1974), generalized random forest (Athey et al., 2019), Shapley values (Shapley, 1988)	Impact heterogeneity of AES participation across several environmental domains, many insignificant effects and even adverse effects; targeting specific farm cohorts might lead to increased scheme effectiveness
Empirical study 4/ chapter 6	Agroforestry and wood-based land-use systems adoption in response to regional weather extremes	Bavarian farms cultivating crops, primary data obtained through survey in 2020; geo-spatial weather data	Discrete choice experiment, correlated random parameter multinomial logit model, extreme weather simulation (Ramsey et al., 2021)	Agroforestry might have a very high probability of being adopted in the medium to long-term, especially after long-lasting extreme weather periods

Zealand (Soliman & Djanibekov, 2021), China (e.g. Song & Chen, 2019; Yang et al., 2021), Pakistan (Ullah et al., 2016) and Chile (Angulo-Meza et al., 2019).

Results on the current state of eco-efficiency in agriculture are inconclusive. Several studies found high (e.g. Godoy-Durán et al., 2017; Stępień et al., 2021), modest (e.g. Bonfiglio et al., 2017; Orea & Wall, 2017; Soliman & Djanibekov, 2021; Stępień et al., 2021) as well as rather low levels (e.g. Picazo-Tadeo et al., 2012, 2011; Pishgar-Komleh et al., 2020) of eco-efficiency. However, studies that regarded eco-efficiency over time, mostly found an improvement in eco-efficiency scores over time (e.g. Staniszewski, 2018; Yang et al., 2021)

While most eco-efficiency studies have focused on the farm-level, there exists also a variety of studies at the regional (Coluccia et al., 2020; Grzelak et al., 2019; Song & Chen, 2019) and country-level (Pishgar-Komleh et al., 2021; Staniszewski, 2018). There is also a wide range of ecological pressures considered in agricultural eco-efficiency studies, including fertilizer and pesticide damages (e.g. Bonfiglio et al., 2017), energy consumption (e.g. Gadanakis et al., 2015), water usage (e.g. Song & Chen, 2019), land consumption (e.g. Grzelak et al., 2019), soil degradation (Eder et al., 2021), greenhouse gas emissions (e.g. Martinsson & Hansson, 2021), waste management (e.g. Godoy-Durán et al., 2017) and biodiversity loss (Beltrán-Esteve et al., 2014).

Numerous studies have also assessed potential eco-efficiency drivers. Beltrán-Esteve et al. (2014), Eder et al. (2021), Picazo-Tadeo et al. (2011), and Gadanakis et al. (2015) showed that high levels of technical efficiency are positively correlated with high eco-efficiency scores. What is more, there seems to be a positive relationship between agri-environmental programs and farms' eco-efficiency (Bonfiglio et al., 2017; Gadanakis et al., 2015; Picazo-Tadeo et al., 2011). Future prospects of generational renewal are negatively associated with eco-efficiency (Bonfiglio et al., 2017; Pérez Urdiales et al., 2016). However, with respect to other potential drivers such as farm size, farm manager's age, education, topography, agricultural training, results are not unambiguous (Bonfiglio et al., 2017; Gadanakis et al., 2015; Godoy-Durán et al., 2017; Pérez Urdiales et al., 2016; Picazo-Tadeo et al., 2011; Soliman & Djanibekov, 2021; Stępień et al., 2021).

Almost all mentioned studies are exclusively based on DEA. Nevertheless, several methodological advancements have been suggested in the literature. For instance, Picazo-Tadeo et al. (2012) use a directional distance function to asses the eco-efficiency of Spanish olive-growers. Beltrán-Esteve et al. (2014) extend this approach and make use of the metafrontier approach proposed by O'Donnell et al. (2008). Furthermore, Kortelainen (2008) embedded the concept of eco-efficiency into a dynamic setting allowing for comparisons over time. In an attempt to transfer the eco-efficiency concept to a SFA setting in order to accommodate random noise and allow for substitutability between environmental pressure, Orea & Wall (2017) study the eco-efficiency of Spanish dairy farms. Their findings are strikingly consistent with the DEA results in Pérez Urdiales et al. (2016), confirming that SFA is an appropriate method for estimating eco-efficiency

Furthermore, there is also a variety of alternative models to incorporate and compare pollution in production technologies, e.g., environmentally-adjusted production efficiency models (Färe et al., 1986), material balance principle-adjusted models (Førsund, 2018), and multiple equation environmentally-adjusted efficiency models (Murty et al., 2012, see also Dakpo et al. (2020) for an extensive overview).

For instance, Skevas et al. (2018d) combine farm accountancy data from a large sample of dairy farms in the Netherlands with data on nutrient surpluses to estimate the impact of farm intensification on environmental efficiency using a hyperbolic distance function. Ait Sidhoum et al. (2022) analyzed trade-offs between economic, environmental and social sustainability in Spanish crop farming using a stochastic frontier latent-class approach. Serra et al. (2014), Malikov et al. (2018) and Lamkowsky et al. (2021), and Tsagris & Tzouvelekas (2022) analyzed the relationship between farm production activities and nitrogen pollution. Ait Sidhoum et al. (2019) additionally evaluated pesticide pollution. Eder (2022) assessed the trade-offs between marketed agricultural outputs and soil erosion. Finally, Dakpo et al. (2017) and Dakpo & Lansink (2019) used by-production models and considered emissions of greenhouse gases as polluting output. Basically, all of these studies conclude that there is a lot of potential for farms to further decrease environmental pollution.

1.3.2 Ecological effects of agri-environmental schemes in Europe

Volunatary AES are part of the common agricultural policy (CAP), which is the main policy framework guiding agricultural legislation (Hasler et al., 2022). More generally, AES are regarded as part of the wider payments for environmental services regime (Wunder et al., 2020). They are aimed at reducing adverse environmental impacts of farming activities by promoting more environmentally-friendly practices (Hasler et al., 2022). There are several overview articles on AES, e.g., Zimmermann & Britz (2016) provide an overview of AES participation and its drivers. Schomers & Matzdorf (2013) summarize the literature on scheme design and Hasler et al. (2022) give a general overview of the European agri-environmental policy.

AES in the context of the CAP have shown mixed success across Europe in terms of meeting environmental targets. Depending on the specific AES and the indicators under investigation, they have been found to be either beneficial (Batáry et al., 2015; Bright et al., 2015; Dadam & Siriwardena, 2019; Dal Ferro et al., 2016; MacDonald et al., 2012; Wuepper & Huber, 2021), ineffective (Bartolini et al., 2021; Bellebaum & Koffijberg, 2018; Calvi et al., 2018; Granlund et al., 2005; Kaligarič et al., 2019; Kleijn et al., 2004), or even detrimental (Baer et al., 2009). The environmental effectiveness of AES has intensively been studied across the continent, e.g. in Italy (e.g. Bartolini et al., 2021; Bertoni et al., 2020; Gatto et al., 2019), Germany (e.g. Pufahl & Weiss, 2009; Uehleke et al., 2022), France (e.g. Chabé-Ferret & Subervie, 2013; Kuhfuss & Subervie, 2018), Switzerland (e.g. Mack et al., 2020; Wuepper & Huber, 2021), and Denmark (Mahmoud & Hutchings, 2020). Arata & Sckokai (2016) compared AES performance across five European countries (UK, Spain, France, Germany, and Italy) and found heterogeneous country-specific effects. Furthermore, several studies conducted additional cost-benefit analyses, which found limited cost-effectiveness of most AES (see Bertoni et al., 2021, 2020; Chabé-Ferret & Subervie, 2013; Faria & Morales, 2020; Gómez-Limón et al., 2019a).

Most recent econometric impact assessments of AES have statistically been based on some sort of matching algorithm (e.g., propensity score matching or coarsened exact matching) in combination with the difference-in-difference estimation approach (e.g. Bertoni et al., 2020; Chabé-Ferret & Subervie, 2013; Uehleke et al., 2022). While these studies have typically been based on farm-level data, there have also been AES evaluations at the field (Mahmoud & Hutchings, 2020), landscape (Tanner & Fuhlendorf, 2018), and regional (Dumangane et al., 2021) level.

The effects of AES have been analyzed across a large variety of environmental domains and indicators, such as water quality (Mahmoud & Hutchings, 2020), crop diversification (Bertoni et al., 2020), planting of cover crops (Chabé-Ferret & Subervie, 2013), organic farming (Mahmoud & Hutchings, 2020), GHG emissions (Bertoni et al., 2021), input use (Kuhfuss & Subervie, 2018), biodiversity (Bougherara et al., 2021), bird conservation (Faria & Morales, 2020), grassland maintenance (Uehleke et al., 2022), and biodiversity conservation areas (Wuepper & Huber, 2021)[1].

The overall environmental effectiveness of AES also depends on scheme uptake (Hasler et al., 2022). There is a stream of literature analyzing farmers' decision to take part in such schemes by linking their preferences to scheme-specific characteristics and/or to farmer and farm-specific features (e.g. Barghusen et al., 2021; Defrancesco et al., 2018; Del Rossi et al., 2021; Leonhardt et al., 2022; McGurk et al., 2020; Zimmermann & Britz, 2016).

Furthermore, the question of how to adjust the design of AES to improve the delivery of a wide range of ecosystem services has been studied intensively (see e.g. Armsworth et al., 2012; Birge et al., 2017; Burton & Schwarz, 2013; Fuentes-Montemayor et al., 2011; Kuhfuss et al., 2016; Latacz-Lohmann & Breustedt, 2019; Latacz-Lohmann & Van der Hamsvoort, 1997; Westerink et al., 2017, 2014). Prominent suggestions for scheme adaptations to trigger a higher return of ecosystem services include the following: defining clear environmental objectives and specifying the relationship between environmental pressures and AES (European Court of Auditors, 2011), linking payments to the actual delivery of desired outcomes (i.e. results based) rather than to management actions (i.e. action/input based) (Birge et al., 2017; Burton & Schwarz, 2013), focusing on (cooperative) landscape management contracts instead of pursuing contracts for individual land management units (Fuentes-Montemayor et al., 2011; Westerink et al., 2017, 2014), introducing a collective bonus (Kuhfuss et al., 2016), setting up auctions (Vergamini et al., 2020) or avoiding standardized payment rates, which ignore the heterogeneity of compliance costs across farmers (Armsworth et al., 2012; Latacz-Lohmann & Breustedt, 2019; Latacz-Lohmann & Van der Hamsvoort, 1997). Until today, there have been almost exclusively action-based AES (i.e. farmers adopt promoted management practices) available, which is why there is very little ex-post empirical evidence on varying payment scheme designs. Recently, Wuepper & Huber (2021) compared result-based and action-based AES in Switzerland and found result-based payments to have a positive environmental effect.

[1] The impacts of AES have also been analyzed with respect to non-environmental outcomes like TFP (Baráth et al., 2020; Mennig & Sauer, 2020), farm survival (Lovén & Nordin, 2020) or regional employment (Dumangane et al., 2021)

An aspect that has often been neglected is the improvement of targeting to increase AES effectiveness. Authors such as van der Horst (2007), Langpap et al. (2008), Desjeux et al. (2015), Früh-Müller et al. (2019), Uthes et al. (2010) and Perkins et al. (2011) stress the importance of spatial targeting and adjusting agri-environmental measures according to land use. It has been shown that both effectiveness and efficiency of AES increase if payments are well-tailored and well-targeted in space and time (Armsworth et al., 2012; Pe'er et al., 2020; Wätzold et al., 2016). However, many of these studies are not based on rigorous impact assessments and have a more correlational character, or rely on simulation approaches (e.g. Longo et al., 2021; Mahmoud & Hutchings, 2020).

1.3.3 Agroforestry, short rotation coppice adoption and climate change

Agroforestry and other wood-based land-use systems such as short rotation coppice (SRC) are recognized to play an important role in synergistically approaching adaptation and mitigation (see e.g. Cardinael et al., 2021; Duguma et al., 2014; Schoeneberger et al., 2012; van Noordwijk et al., 2014, 2011; Verchot et al., 2007).

The elicitation of farmers' preferences for agroforestry and woody perennials has been the subject of multiple studies, including Gillich et al. (2019) and Pröbstl-Haider et al. (2016), who analyzed farmers' preferences for SRC in Germany and Austria using discrete choice experiments. Mercer & Snook (2005), Schaafsma et al. (2019), and Oviedo et al. (2021) used discrete choice valuation methods to elicit farmers' preferences for agroforestry in Mexico, Malawi, and Spain.

More commonly, farmers' adoption of trees on agricultural land has been analyzed ex-post by linking farm and household characteristics to the adoption of such systems using econometric binary outcome models. Amare & Darr (2020) provide an overview of such analyses and find that a total of 151 variables have been used to study agroforestry adoption. Many of these studies focus on the Global South, e.g. Amusa & Simonyan (2018), Adesina & Chianu (2002), Beyene et al. (2019), and De Giusti et al. (2019). Classical adoption drivers are for instance age, education and income (see e.g. Pattanayak et al., 2003). Additionally, psychological and sociological factors, such as attitudes, subjective norms, self-efficacy or perceived behavioral control have been brought into focus (Buyinza et al., 2020; McGinty et al., 2008; Meijer et al., 2015). Various wood-based land use systems have been considered in such adoption studies, including SRC, alley cropping, coffee-integrated agroforestry, fruit trees, hedgerows and many more (Adesina & Chianu, 2002; Bannister & Nair, 2003; Mfitumukiza et al., 2017). Furthermore, some studies have specifically focused on (potential) barriers and constraints regarding the adoption of agroforestry, e.g. Djalilov et al. (2016), Mattia et al. (2018), and Sollen-Norrlin et al. (2020).

Other studies have used investment analysis to evaluate the economic potential of agroforestry and other tree-based land-use systems (e.g. Frey et al., 2013; Lasch et al., 2010). Another important stream of the literature has conducted adoption assessments by means of land-use simulations. Gosling et al. (2021) used modern portfolio theory and robust multi-objective optimization to assess the economic potential of different agroforestry systems in Panama. Gosling et al. (2020) and Gosling et al. (2020) used this approach under consideration of ten different adoption drivers also accounting for ecological goals and farmers' preferences. Romanova et al. (2021) focused on the temporal dynamics of agroforestry adoption process (using a qualitative research approach), an important aspect that had often been neglected in the literature.

When it comes to assessing farmers' preferences and adoption decisions in the face of climate change and (extreme) weather, there are only a few studies. Lasch et al. (2010) and Gomes et al. (2020) projected the cultivation potential for SRC in eastern Germany and coffee-agroforestry in Brazil, taking into account various climate change scenarios until 2050. Schaafsma et al. (2019) discuss agroforestry in the context of climate-smart agriculture in Malawi. Lasco et al. (2016) discussed the role of agroforestry given farmers' perceptions of climate change. Moreover, Paul et al. (2017) use robust optimization to simulate the adoption of various tree-based land-use systems under consideration of climate uncertainty. The only study explicitly evaluating the effect of weather on agroforestry adoption was conducted by Mfitumukiza et al. (2017) using a case study in Uganda.

1.4 Short description of the case study region: Bavaria

As mentioned above, the empirical studies presented in Part II of this thesis focus on farming in Bavaria. Located in the southeast of Germany, Bavaria belongs to the core regions of agricultural production within the European Union (EU). It comprises seven regional districts, 71 counties and 2056 municipalities. In 2019, there were a total of 105,297 farms, which managed 3.1 million hectares of land (44% of total land). Crop land accounted for 65%, and grassland made up 35%.

233 thousand people were employed in the agricultural sector in Bavaria in 2016 (StMELF, 2021).

In terms of natural conditions, farming takes place along an elevational gradient of 1500 m (from 100 m in Northwest Bavaria to 1600 m in Southeast Bavaria) and a macroclimatic gradient with a mean annual temperature range between 3 °C and 10 °C and an annual precipitation of 470–1592 mm (from 1960 to 2020). Its natural conditions range from pre-alpine and alpine areas with high precipitation and rather clayey, limestone and dolomite-based soils to regions with flat land and fertile loess soils to dry, marlstone, limestone and dolomite based hillside locations. Figure 1.3 provides an overview of Bavaria's heterogeneous natural conditions.

These conditions are well-suited for, and reflected by, various agricultural production systems such as crop farming, intensive and extensive dairy farming, pig and cattle fattening and breeding, poultry farming, vegetable farming, orcharding, hop production and viticulture. This heterogeneity of farming systems represents to some extent the European agricultural sector. For instance, farms in Bavaria managed on average 34.7 ha of land in 2014 which is similar to average farm sizes in, e.g., Ireland, Belgium and the Netherlands (European Statistical Office, 2020). Furthermore, Bavaria is one of the largest milk-producing regions in the EU (Frick & Sauer, 2018), and the Bavarian farm labor structure and livestock count can be seen as representative for a large number of European farms in that average numbers are close to European averages (European Statistical Office, 2020). Bavarian dairy farmers kept on average 34.1 livestock units (LUs) of dairy cows between 2005 and 2014. This value lies only slightly above the average for all European regions (30.4). On average, Bavarian pig farms managed 154.4 LUs of pigs, which was approximately equal to the European average of 154.8 LUs. As for the labor structure, on average 1.6 average work units (AWUs) worked on Bavarian farms between 2005 and 2014, while the European average was 1.5 AWUs in the same period (European Statistical Office, 2020).

1.5 Outline of the thesis

The remainder of this thesis is structured as follows. Chapter 2 provides a general overview of applied concepts and methods of this thesis. Part II consists of the following four empirical studies focusing on the nexus of agricultural production and environmental change in Bavaria:

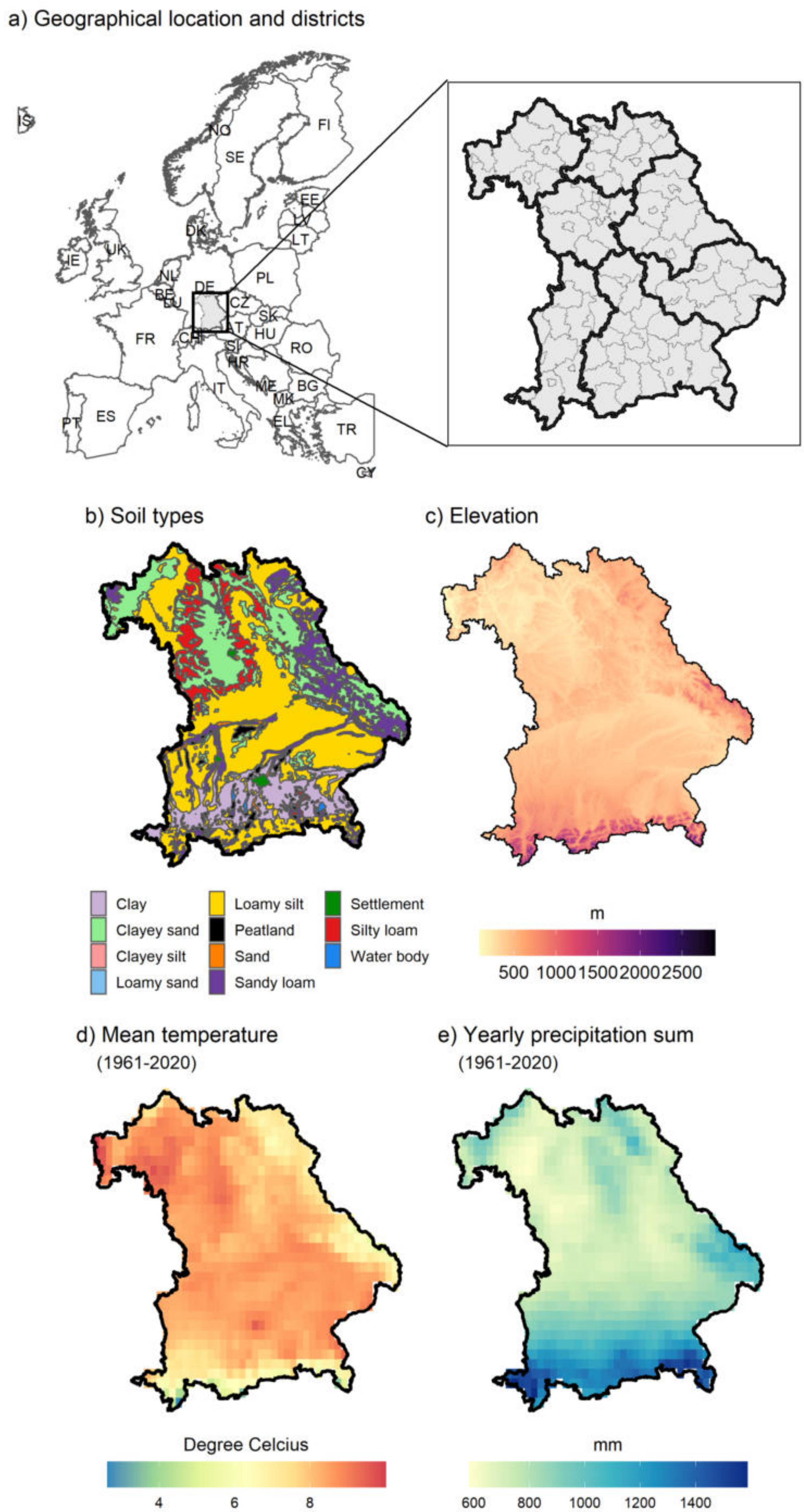


Figure 1.3: Description of the case study region, Bavaria. Sources: BGR (2022); Cornes et al. (2018).

- Chapter 3: Study I – Greenhouse gas emissions and eco-performance at farm level – a parametric approach

  Stetter, C., & Sauer, J.(2022). Greenhouse Gas Emissions and Eco-Performance at Farm Level: A Parametric Approach. *Environmental and Resource Economics*, 81, 617–647. DOI: `https://doi.org/10.1007/s10640-021-00642-1`

- Chapter 4: Study II – Are intensive farms more emission-efficient? Evidence from German dairy farms

  Stetter, C., Wimmer, S. & Sauer, J.(2022). Are Intensive Farms More Emission-Efficient? Evidence From German Dairy Farms. *Journal of Agricultural and Resource Economics*. DOI: `doi:10.22004/ag.econ.31675`

- Chapter 5: Study III – Using machine learning to identify heterogeneous impacts of agri-environment schemes in the EU: A case study

  Stetter, C., Mennig, P. & Sauer, J.(2022). Using Machine Learning to Identify Heterogeneous Impacts of Agri-Environment Schemes in the EU: A Case Study. *European Review Of Agricultural Economics*. DOI: `https://doi.org/10.1093`

- Chapter 6: Study IV – Tackling climate change: Agroforestry adoption in the face of regional weather extremes

  Stetter, C., & Sauer, J.(2022). Tackling Climate Change: Agroforestry Adoption in the Face of Regional Weather Extremes. *Working paper*.

Finally, Part III summarizes the main findings and gives a general discussion of the presented research articles, provides policy implications as well as recommendations for future research.

2 An overview of applied concepts and methods

This chapter provides an overview of the theoretical background of this thesis and a short glimpse into the methods applied in Part II. Conceptually, this thesis combines microeconomic theory with state-of-the art econometric methods and real-world farm-level data.

2.1 Production analysis

The main theoretical focus of this thesis is on applied production economics, which represents the theoretical basis for studies I to III. Agricultural production economics focuses on the producers of agricultural commodities, their goals and objectives, their choices in terms of resource allocation, input and output quantities and their (economic) environment (Debertin, 2012).

2.1.1 The production technology and technical efficiency

In production economics, all production processes are regarded as a transformation of inputs (e.g. materials, labor, and capital) to produce outputs (goods and services) (Kumbhakar et al., 2015). This basic principle is applicable to a large array of different economic entities such as factories, firms, and farms or to non-profit organizations like schools or hospitals, but also to smaller units within these entities (e.g., bank branches, retail stores, agricultural enterprises) and to the macro-level like regions, countries or sectors (Coelli & Rao, 2005b).

Mathematically, the transformation of a vector of inputs (x) into a vector of outputs (q) can be described by means of a production technology set, S and reflects the technological production possibilities of a firm (Coelli & Rao, 2005b):

$$S = \{(x, q) : \text{x can produce q}\}. \tag{2.1}$$

For the case of a single output scalar q, the production possibilities of a firm can conveniently be described by a production function (Kumbhakar et al., 2015):

$$q = f(x_1, x_2, \ldots, x_J) \equiv f(x) \tag{2.2}$$

The production function $f(x)$ reflects the maximum attainable output q for a given available vector of inputs x. In production theory, it is usually assumed that production functions are consistent with a set of axiomatic properties (see Chambers, 1988, p.9 for a detailed discussion on this).

For now, it has implicitly been supposed that all production activities are on the frontier of the technology set and obtain maximum output (Kumbhakar et al., 2015), i.e. farms are considered technically efficient. However, a firm is technically inefficient with respect to output if it could produce a higher level of output for the given inputs (output-oriented inefficiency). Following Shepherd (1970), technical efficiency (TE) can be defined as

$$TE = \frac{q}{q^*} \Leftrightarrow q = q^* \cdot TE \quad 0 \leq TE \leq 1, \tag{2.3}$$

where q is the observed output quantity and q^* is the maximum attainable output quantity with the observed input quantities x. The concept of technical inefficiency is graphically demonstrated in Figure 2.1. Production function $f(x)$ defines the maximum attainable output q^* for a given input level x. Point A is technically inefficient because it could produce more output at the current level of x. The distance $\overline{AB}$ reflects the output loss due to technical inefficiency (Kumbhakar et al., 2015). Figure 2.1 equivalently demonstrates the case for input-oriented technical inefficiency for point A in that it could produce the same amount of output with less input, reflected by the distance $\overline{AC}$.

2.1.2 Stochastic frontier analysis

There exists a variety of parametric and non-parametric approaches to obtain empirical estimates of the production technology and associated TEs based on observed production data. One of the most prominent and widely applied method is the stochastic frontier analysis (SFA), a parametric statistical analysis technique (Kumbhakar & Lovell, 2000). The concepts presented above are deterministic, i.e. the difference between observed output and maximum attainable output is solely attributed to technical inefficiency (Kumbhakar & Lovell, 2000), and neglect statistical noise (e.g. random shocks outside producers' control like weather) and

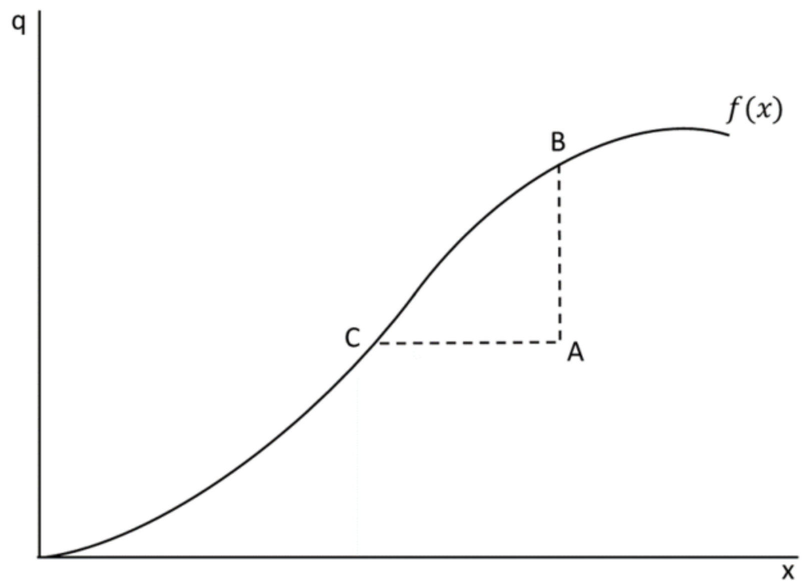


Figure 2.1: Illustration of technical inefficiency for production frontier $f(x)$. Point A is inefficient with respect to output q because it could produce more q using the same amount of input x, which is reflected by the distance $\overline{AB}$. Point A is also inefficient with respect to input x because it could use less input x to produce the same level of q, which is reflected by distance the $\overline{AC}$.

are very sensitive to positive outliers (Henningsen, 2019). SFA is able to account for this stochasticity in the production process. Consequently, one can reformulate Equation 2.3 by defining $TE = exp(-u)$ and adding a producer-specific random shock term $exp(\nu)$:

$$y = f(x) \cdot \exp(-u) \cdot \exp(\nu). \tag{2.4}$$

To facilitate estimation and following the seminal papers of Aigner et al. (1977) and Meeusen & van den Broeck (1977), (2.4) can be represented in logarithmic form by

$$\ln y = \ln f(x) - u + \nu \quad u \geq 0, \tag{2.5}$$

where ν accounts for idiosyncratic errors in the estimation and u is a positive, one-sided error term accounting for technical inefficiency. Hence, in the stochastic frontier model, there is a composite error term ($\epsilon = \nu - u$).

The stochastic frontier model is commonly estimated by way of maximum likelihood estimation (Kumbhakar & Lovell, 2000). This estimation technique requires assumptions on the distribution of the two error terms u and ν. Usually, the noise term ν is assumed to be normally distributed with zero mean and constant variance σ_ν^2. The distributional assumption for the inefficiency term u varies but often follows a positive half-normal distribution with constant scale parameter σ_u^2 (see e.g. Henningsen, 2019):

$$\nu \sim N(0, \sigma_\nu^2) \tag{2.6}$$

$$u \sim N^+(0, \sigma_u^2) \tag{2.7}$$

Finally, the researcher must choose a functional form for $f(x)$. Common choices are the Cobb-Douglas, quadratic or the translog functions, respectively (Coelli & Rao, 2005b).

2.1.3 Environmental pressure generation and eco-efficiency

The above-described production technology does not account for potential pollution stemming from production processes. As stated in the beginning, current agricultural practices are associated with multiple environmental problems (see e.g. Campbell et al., 2017). Ignoring these environmental damages would yield an incomplete picture of the impacts of agricultural production activities.

A variety of models have been proposed in the literature to incorporate environmental pollution in production technologies: environmentally-adjusted production efficiency models, material balance principle-adjusted models, and multiple equation environmentally-adjusted efficiency models (Dakpo et al., 2020; Lauwers, 2009). The latter category includes the by-production model proposed by Murty et al. (2012), which defines the global technology as the intersection of two subtechnologies – one for good outputs (e.g., wheat or milk) and one for bad outputs (e.g., nitrogen leaching or GHG emissions).

Another concept developed by Kuosmanen & Kortelainen (2005) is eco-efficiency, which is the theoretical basis for the studies in chapters 3 and 4. It assesses economic activity from an environmental impact perspective without direct recourse to physical inputs and outputs. Starting point for this concept is the pressure-generating technology set ($PGTS$) (Kortelainen, 2008; Kuosmanen & Kortelainen, 2005; Picazo-Tadeo et al., 2012), which describes how ecological pressures s translate to economic returns y:

$$PGTS = \Big[(s, y) : \text{ economic returns } y \text{ can be generated with ecological pressures } s\Big]. \tag{2.8}$$

It contains all technically and economically feasible combinations of economic output (y) and environmental pressures (s).

Based on the $PGTS$, eco-efficiency evaluates the ability of firms to generate a higher level of economic returns at a given level of environmental damage, or conversely, to generate a given level of economic returns with less environmental damage. Figure 2.2 illustrates this concept, which is very similar to the traditional production technology and TE concept presented above.

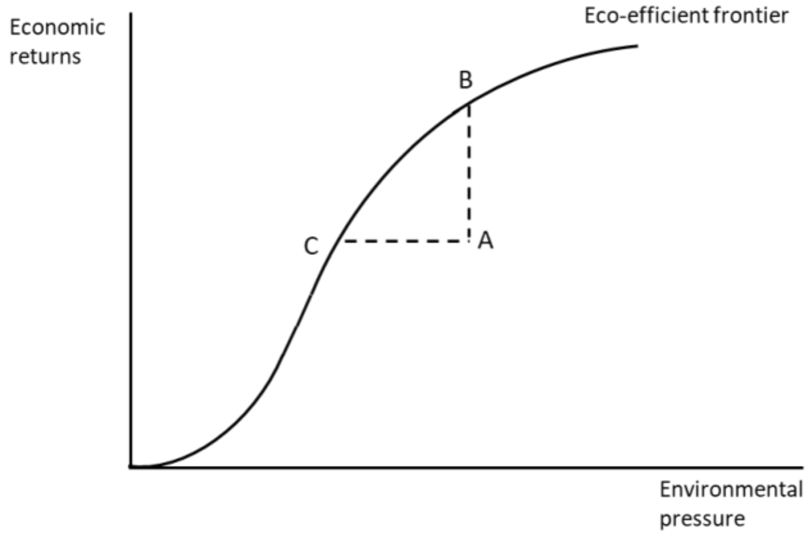


Figure 2.2: Illustration of eco-inefficiency. Point A is eco-inefficient with respect to economic returns because it could generate higher economic returns with the same amount of environmental pressure, which is reflected by the distance $\overline{AB}$. Point A is also eco-inefficient with respect to environmental pressure because it could cause less environmental pressure to generate the same level of economic returns, which is reflected by distance the $\overline{AC}$.

Usually, eco-efficiency is estimated using mathematical programming methods, i.e. DEA. However, this ignores the stochastic nature of the relationship between economic returns and environmental pressures. Study I and Study II make use of the fact that most concepts underlying eco-efficiency are very similar to traditional production economic concept and assess farms' ecological-environmental performance with respect to greenhouse gas emissions using SFA methods.

Although the eco-efficiency framework might axiomatically not be the most accurate approach, eco-efficiency scores are easy to interpret and express the environmental pressure mitigation potential in a tangible number, thus making them a meaningful index allowing for policy-relevant conclusions regarding sustainable management and efficient use of natural resources in the agricultural sector (Coluccia et al., 2020).

2.1.4 The production technology and action-based agri-environmental schemes

Turning back to classical production analysis, one can also look at production from a multiple output perspective. In that case, the technology set (2.1) can be represented by an output set reflecting production possibilities for a given available set of inputs (Coelli & Rao, 2005b):

$$P(x) = \{q : \text{x can produce q}\} = \{q : (x, q) \in S\}. \tag{2.9}$$

These outputs can be both marketed (e.g. wheat or milk) and non-marketed outputs (e.g. GHG emissions or nutrient leaching). Given the action-based nature of most AES, i.e. participation limits their production possibilities, (2.9) can be re-expressed as:

$$\begin{aligned} P(x) &= \{q : x \text{ can produce } q \text{ for a given AES participation status } w\} \qquad (2.10) \\ &= \{q : (x, q, w) \in S\}. \end{aligned}$$

Figure 2.3 graphically illustrates the case of one marketed and one non-marketed output. In this stylized representation, it is assumed that farmers decide either to participate or not to participate in an AES. Accordingly, farms face two potential production possibility curves (PPCs)[1], where participation in an AES causes an inward shift of the PPC. Based on this model, Chapter 5 seeks to answer the question how participation in AES affects the environmental outcome of farming activities across multiple environmental domains.

The fact that one can only observe the outcome of one participation status at a time (either a farm takes part or not) represents the fundamental problem of causal inference, which is addressed by the potential outcomes framework developed by Neyman (1923) and Rubin (1974) (see next section).

2.2 Potential outcomes, conditional average treatment effects and causal machine learning

Within the potential outcomes framework, a causal effect is defined as the difference between two states of the world (Cunningham, 2021). In the context of this

[1]The PPC graphically depicts all possible combinations of two outputs that could potentially be produced at a given input bundle. Another term for PPC is production possibility frontier (PPF), which can be used interchangeably.

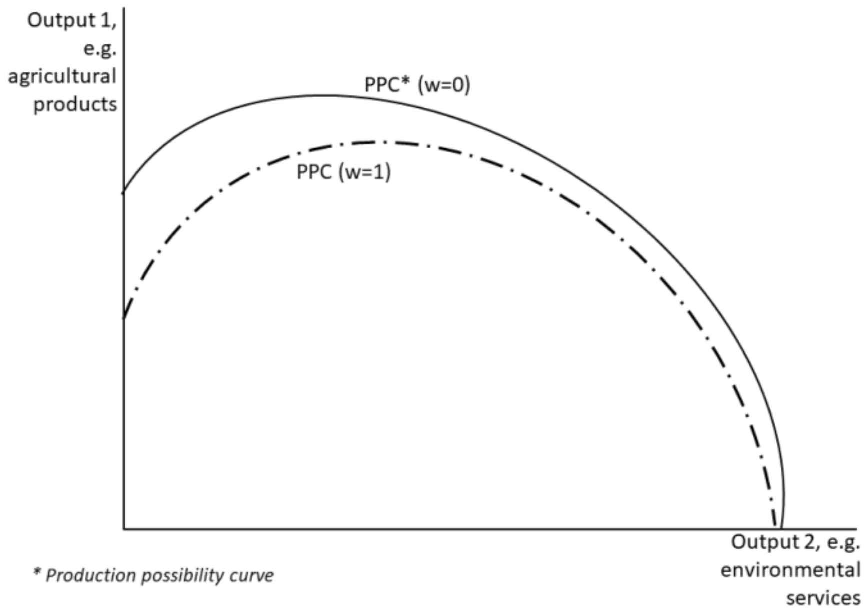


Figure 2.3: Illustration of the potential production possibility curves when not participating in an action-based agri-environmental scheme PPC($w = 0$) and when participating in an action-based agri-environmental scheme PPC($w = 1$). All output combinations on and below a production possibility curve are economically feasible. Hence, participation in an action-based agri-environmental scheme limits a farm's production possibilities.

thesis, this would mean comparing the environmental performance of a farm in a world, in which it takes part in an AES (Y_i^1) versus a world in which it does not take part in an AES (Y_i^0). The problem is that one cannot observe both outcomes for one farm. The unobservable outcome is called counterfactual. The treatment effect τ_i for individual i is defined as

$$\tau_i = Y_i^1 - Y_i^0. \tag{2.11}$$

It is not possible to estimate (2.11) because the counterfactual outcome is unknown. However, it is possible to estimate an average treatment effect (ATE) when individuals are randomly selected and assigned into a program by utilizing the expected values of the potential outcomes (Cuong, 2009):

$$ATE = \mathbb{E}\Big[\tau_i\Big] = \mathbb{E}\Big[Y_i^1 - Y_i^0\Big] = \mathbb{E}\Big[Y_i^0\Big] - \mathbb{E}\Big[Y_i^1\Big]. \tag{2.12}$$

However, there are two major problems associated with this approach. First, as in many other settings, participation in AES is not random. Instead, farmers choose (self-select) whether or not they participate in a program. Hence, the calculated treatment effect in (2.12), i.e. the simple difference between the expected values

$\mathbb{E}\Big[Y_i^0\Big]$ and $\mathbb{E}\Big[Y_i^1\Big]$, is likely not solely attributable to the difference in participation status D_i (equals 1 if one participates in the program, and 0 otherwise), but is also due to inherent differences between the two groups (e.g. farm characteristics such as age, education, and size). Without further assumptions, it is impossible to identify the effect of a policy given the observational nature of the research problem. Therefore, Study III invokes the conditional independence assumption (Rubin, 1977), i.e. participation status D_i is independent of unobservable features conditional on a set of contextual characteristics X_i: $Y_i^1, Y_i^0 \perp\!\!\!\perp D_i \mid X_i$. Furthermore, common support is assumed to rule out perfect predictability of program participation, i.e. individuals with the same X have a positive probability of being both participants and non-participants: $0 < P(D_i = w \mid X) < 1$.

An often neglected aspect regarding the evaluation of agri-environmental programs (and other policies) is effect heterogeneity. The impacts of AES are expected to vary across farm households depending on their farming context (e.g. factor endowment, natural conditions, etc.) (Pufahl & Weiss, 2009). Although acknowledged by many previous studies on the subject, most of them could only estimate average effects on the basis of traditional statistical methods (e.g. Arata & Sckokai, 2016; Bertoni et al., 2020; Chabé-Ferret & Subervie, 2013). Neglecting effect heterogeneity could potentially lead to flawed policy conclusions. For instance, assume the true ATE of a policy is +10, containing half of the subjects with a treatment effect of +15 and the other half of −5. The positive ATE conceals the fact that this policy has a negative impact on half of the subjects.

Thus, the conceptual approach of Study III is based on the conditional average treatment effect (CATE) that allows to obtain individualized AES effects. Suppose a set of i.i.d. farm households $i = 1, \ldots, n$, for which we observe (X_i, Y_i, D_i), where $X_i = x \in \mathbb{R}^p$ is a vector of p contextual covariates, describing the individual farming context and containing all determinants of Y^0 and Y^1 as well as the determinants of the participation decision. As above, $Y_i \in \mathbb{R}$ is the outcome variable of interest (e.g. an indicator reflecting environmental performance), and $D_i \in \Big\{0, 1\Big\}$ is the policy dummy for participation and non-participation in AES. Given the potential outcomes Y_i^0 and Y_i^1, for each farm i that is (uniquely) characterized by its contextual feature vector x, the CATE can be expressed by:

$$\tau(x) = \mathbb{E}\Big[Y_i^1 - Y_i^0 \mid X_i = x\Big]. \tag{2.13}$$

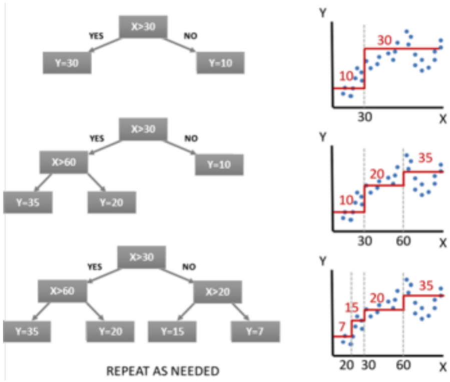
Hence, the set of contextual variables x defines the extent to which the CATE is individualized, i.e. the better the contextual variables (x) define each farm the more accurate become the individualized treatment effect estimates $\tau(x)$. Estimating the CATE with a large set of contextual variables that contains non-linearities and interactions is a non-trivial task (Athey & Imbens, 2019). However, in recent years, a series of novel estimation methods have been introduced combining causal inference and machine learning (ML) to obtain accurate CATE estimates (see e.g. Athey & Imbens, 2016; Athey et al., 2019; Künzel et al., 2019; Wager & Athey, 2018).

ML algorithms have primarily been developed for prediction tasks (Storm et al., 2020).[2] Popular ML prediction tools include shrinkage methods such as the *least absolute shrinkage and selection operator* (LASSO), *Ridge regressions*, *neural networks*, and tree-based methods such as *(boosted) classification and regression trees (CART)* or *random forests (RFs)* (see e.g. James et al., 2021, for an overview).

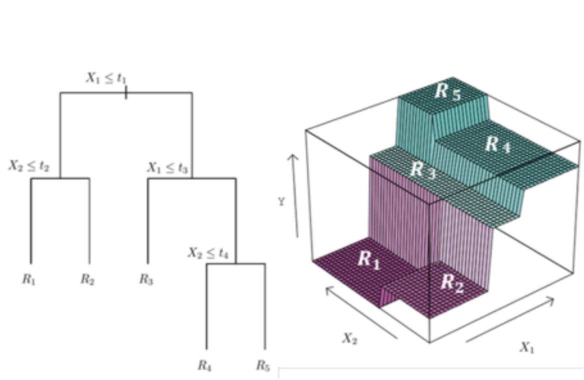
Based on these methods, a series of ML algorithms for the estimation of causal effects have been suggested, e.g. meta-learners (Künzel et al., 2019) or neural networks for causal inference (Farrell et al., 2021). However, CART and RF-based algorithms have been playing a key role in the estimation of CATEs (Storm et al., 2020). Figure 2.4 provides an intuition as to how these algorithms work. Random forests, a concept developed by Breiman (2001), are basically an ensemble of CARTs, which are grown based on recursive partitioning such that the variable space is divided into binary nodes according to an optimality criterion (e.g. many standard regression tree implementations split by minimizing the in-sample prediction error of the node (Breiman et al., 2017)) until the final nodes (aka leaves) contain a number of observations greater than a given minimum. The average outcome of such a leaf is then the prediction for the observations contained in that leaf. Random forests make predictions in the form of an average across predictions $b = 1, \dots B$ of such CARTs, each of which is grown on a training sample, i.e. a random subsample of the data.

Tree-based causal ML methods include, among others, causal trees (Athey & Imbens, 2016), causal forests (Athey et al., 2019; Wager & Athey, 2018), modified causal forests (Lechner, 2019), and orthogonal random forests (Oprescu et al., 2018). All of these algorithms rest on adjusted splitting rules, such that they are

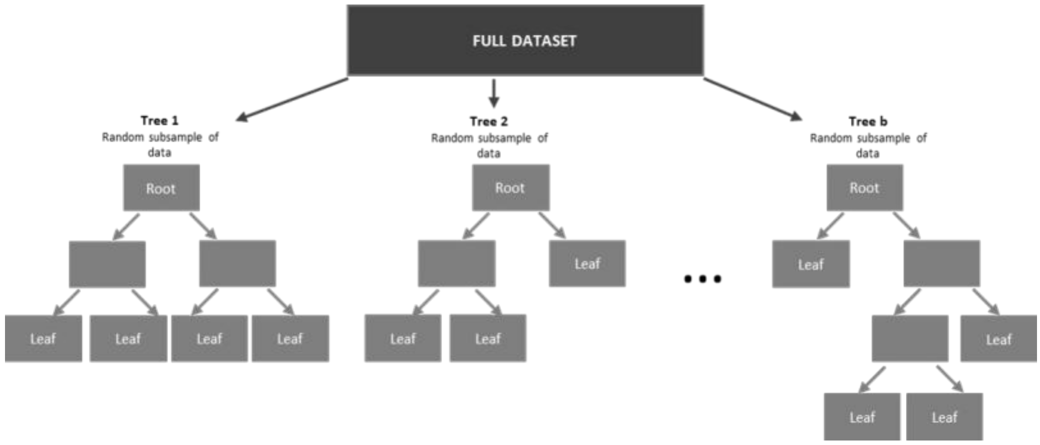
[2] Predicting some output based on a set of explanatory variables is called supervised ML. Another common use of ML are unsupervised approaches, i.e. grouping and clustering of data based in characteristics of observations. (Storm et al., 2020).



(a) Demonstration of a CART with one feature (X) and one response variable (Y). *Source*: Tiffin (2019).



(b) Demonstration of a CART with two features (X_1, X_2) and one response variable (Y) and partition point t. *Source*: James et al. (2021).



(c) Random forests are an ensemble of regression trees based on random subsamples of the full dataset. *Source*:own depiction based on Tiffin (2019).

Figure 2.4: Regression trees and random forest architecture.

optimized for an accurate CATE prediction. The empirical strategy in Study III is based on the causal forest algorithm (Athey et al., 2019; Wager & Athey, 2018), as this is arguably the most widely applied and well-established causal ML algorithm in this context.

A common problem with ML prediction models is the fact that their predictive power and ability to estimate complex models comes at the expense of interpretability (Molnar, 2019; Storm et al., 2020). Based on this problem, *interpretable machine learning* has evolved as a new research discipline. It has put forth a set of model-specific as well as model-agnostic interpretability methods in recent years. Model-agnostic interpretability means that interpretation is separated from estimating/learning a model. This makes the concept very flexible in that it is not bound to any specific ML model (Ribeiro et al., 2016). One such concept that has been enjoying increasing attention are Shapley values (Shapley, 1988). It is the only interpretability concept with a solid theory (Molnar, 2019). Study III uses Shapley values to explore the relationship between the most relevant contextual variables and the impact size of AES.

2.3 Random utility theory and discrete choice experiments

In Study IV, farmers' land-use choices were assessed in response to extreme weather events. To this end, it has been necessary to elicit farmers preferences for different land use alternatives, which is usually done by means of stated preference methods (Louviere et al., 2000). There are several stated preference methods to do that. These methods measure individuals' preferences for alternative choices based on their decisions in hypothetical choice situations (Louviere et al., 2000). Respondents have to state their choice over sets of hypothetical alternatives (Mangham et al., 2009). These alternatives are described by a series of characteristics. Based on the respondents' responses, i.e. choices, the value attributed to each characteristic can be inferred. Usually, individuals' are repeatedly confronted with decision situations, in which the characteristics of the alternatives change (Louviere et al., 2000).

DCEs rest conceptually on random utility maximization (Mariel et al., 2021) following Lancaster' characteristics theory of value (Lancaster, 1966) and random utility theory (McFadden, 1973). Individuals are assumed to maximize their utility. Each individual n obtains a certain level of indirect utility (U) from each choice alternative. In a given decision situation t, they will select alternative i if and only if $U_{it} > U_{jt}, j \neq i$. The indirect utility of an alternative cannot be directly measured, but it can be expressed by a systematic (deterministic) component V, reflecting specific characteristics as well as farmers' individual and location-specific features, plus a random component ϵ, representing unobserved decision-relevant elements (Mariel et al., 2021). Consequently, a farmer n obtains a certain level of indirect utility U_{nit} from a land use alternative i in a choice situation t.

$$U_{nit} = V_{nit} + \epsilon_{nit} \tag{2.14}$$

Furthermore, as agricultural land-use is heavily dependent on weather, it is assumed that farmers' utility also depends on expected weather (c) at the time of the planting decision. Thus, U can be formulated as follows:

$$U_{nit} = f(x_{nit}, c_{nt}; \beta, \gamma) + \epsilon_{nit} \tag{2.15}$$

where β and γ are coefficients to be estimated. Depending on the assumptions about ϵ_{nit}, there are multiple models for estimating the unknown coefficient vectors β and γ, usually based on the (simulated) maximum likelihood method, e.g. the multinomial logit, the mixed logit, and the latent-class logit model (Louviere et al., 2000).

Part II

Empirical Studies

3 Study I – Greenhouse gas emissions and eco-performance at farm level: a parametric approach

Disclaimer:

This version of the article has been accepted for publication, after peer review but is not the Version of Record and does not reflect post-acceptance improvements, or any corrections. The Version of Record (Stetter, C. & Sauer, J. (2022). Greenhouse Gas Emissions and Eco-Performance at Farm Level: A Parametric Approach. Environmental and Resource Economics, 81(3), 617–647) is available online at: `https://doi.org/10.1007/s10640-021-00642-1`.

Acknowledgments

The authors greatly appreciate comments received from all anonymous reviewers who helped to substantially improve earlier versions of the manuscript. Furthermore, the authors would like to thank the participants of the second meeting of the OECD network on agricultural total factor productivity and the environment for their constructive feedback. This research was partially supported by the Bavarian State Ministry of Sciences, Research and the Arts in the context of the Bavarian Climate Research Network (bayklif). Open Access funding enabled and organized by Projekt DEAL.

3.1 Abstract

Agriculture is an important source of greenhouse gas (GHG) emissions and thus contributes considerably to global warming. However, farms can vary substantially in terms of their climatic impact. So far, most policies aiming at reducing GHG emissions from farming have largely been based on findings at the aggregate level, without taking farm heterogeneity properly into account. This study seeks to provide a better understanding of the GHG mitigation potential at the micro-level. We develop a comprehensible analytical framework for analyzing economic-ecological performance by way of stochastic frontier analysis. We introduce the concept of emission efficiency, where we distinguish between persistent and time-varying efficiency. We further analyze farms with respect to their emission-performance dynamics. Results from our (2005–2014) empirical application from Bavaria – an important agricultural region for the EU – show considerable differences in farm-level GHG emissions across different farm types. The same applies to emission efficiencies. Overall, emission performance improved over time. The results have important climate- policy implications as they help to provide better target measures for mitigating GHG emissions from agriculture, without compromising economic performance levels.

3.2 Introduction

Climate change – and its predicted consequences – has been a major topic of interest in research, in politics and the public sphere for several decades. The discussion also highly affects agriculture as it is a driving force behind global warming, accounting for approximately 14% greenhouse gas (GHG) emissions worldwide (European Environment Agency, 2019a). According to the Food and Agriculture Organization of the United Nations (FAO, 2017), climate change is among the biggest challenges agriculture is currently facing in terms of both mitigation and adaptation. The European Union (EU) has acknowledged the impact of farming on climate change and vice versa. Therefore, the European Commission (2019a) has declared climate change mitigation as a specific objective of the future Common Agricultural Policy of the EU (CAP). While farmers are expected to reduce their GHG emissions, the global need for food, fiber and bioenergy from agriculture is steadily rising (FAO, 2017), i.e. agricultural productivity must further increase. This raises the question as to what extent farmers can boost their production while at the same time limiting their release of greenhouse gas emissions into the atmosphere.

In this study, we analyze the performance of farms, giving equal weighting to economic success and to the emission of greenhouse gases. More specifically, our research objective is two-fold. First, we develop a parametric eco-efficiency concept capable of jointly evaluating the ecological and economic performance of businesses over time. Usually, eco-efficiency is analyzed in a nonparametric data envelopment setting. The idea behind the parametrization is to overcome some of the conceptual limitations faced by the model in a nonparametric setting. Second, we seek to approximate GHG emissions for a large sample of German farms by using a novel approach based on the guidelines for national GHG inventories of the Intergovernmental Panel on Climate Change (IPCC, 2006).

Over recent decades, the concept of ecological-economic efficiency, or "eco-efficiency" has gained recognition as a way of evaluating the balance between economic performance and the environmental damage induced by economic activity. Starting with the article by Kuosmanen & Kortelainen (2005), who operationalized the eco-efficiency concept to be calculated by means of data envelopment analysis (DEA), numerous studies in several fields followed, both at the micro- and at the macro-level. Literature focusing on the trade-off between greenhouse gas emissions and economic performance is sparse. Camarero et al. (2014) and Gómez-Calvet et al. (2016) analyze this relationship for European countries. Multiple authors have utilized the above-mentioned eco-efficiency concept in the farming context, ranging from livestock and arable farming to olive growing and horticulture (Bonfiglio et al., 2017; Gadanakis et al., 2015; Godoy-Durán et al., 2017; Pérez Urdiales et al., 2016; Picazo-Tadeo et al., 2012, 2011). Several methodological advancements have been suggested in the literature. For instance, Picazo-Tadeo et al. (2012) use a directional distance function to asses the eco-efficiency of Spanish olive-growers. Beltrán-Esteve et al. (2014) extend this approach and make use of the metafrontier approach proposed by O'Donnell et al. (2008). Furthermore, Kortelainen (2008) embedded the concept of eco-efficiency into a dynamic setting allowing for comparisons over time. In an attempt to transfer the eco-efficiency concept to a parametric setting, Orea & Wall (2017) study the eco-efficiency of Spanish dairy farms in a stochastic frontier setting. Many of the aforementioned studies – and micro-level studies particularly – are rather limited in terms of their scope. For instance, none of the mentioned micro-studies on agriculture have implemented a dynamic approach yet. Also, although having acknowledged the possibility of variable returns to the scale of environmental pressures, very few studies account for it (Bonfiglio et al., 2017). What is more, none of the studies have implemented the possibility of a non-linear structure when building composite indicators of environmental pressures. According to Beltrán-Esteve et al. (2014), one reason for the limited scope of many works on eco-efficiency is the lack of appropriate environmental data at the micro-level. This is also true for GHG emissions. Farms can vary substantially regarding their climatic impact. So far, most policies aiming at reducing GHG emissions from farming have largely been based on findings at the aggregate level, without taking farm heterogeneity properly into account. Recent attempts have been made to obtain GHG data at farm level. Coderoni & Esposti (2014) and Baldoni et al. (2017) present a methodology to gather GHG emissions primarily based on farm accountancy data. Further applications of this method can be found in Baldoni et al. (2018) and Coderoni & Esposti (2018).

This study contributes to the literature by first conceptually extending the eco-efficiency approach based on Kuosmanen & Kortelainen (2005) and Orea & Wall (2017) to a more general stochastic frontier setting. By parameterizing the concept, we allow for variable returns to environmental pressure scale, a nonlinear structure of composite ecological damage, and for inter-temporal comparisons. Furthermore, we decompose eco-performance dynamics into ecological-technical change, eco-efficiency change and scale change, based on a generalized Malmquist productivity index (Orea, 2002). We are able to show that our theoretical approach is equally valid for micro- and macro-level analyses. In our empirical case study, we construct a unique panel dataset and demonstrate how to appropriately approximate GHG emissions at the micro-level, based on multiple data sources for farms in the German federal state of Bavaria. Finally, by applying a state-of-the-art stochastic frontier model, we are able to further distinguish between permanent and time-varying eco-efficiency (Kumbhakar et al., 2014). We show how to assess farms' *emission performance*, i.e. their capability to produce goods and services while causing minimal climatic stress. By relating eco-performance to greenhouse gas emissions, it is possible to detect the relative GHG mitigation potential of individual farms. Such robust evidence is essential for managers and policy-makers that aiming at further optimizing their economic-ecological performance by reducing the release of atmospheric GHGs.

The remainder of this article is organized as follows. In Section 3.3, we develop the conceptual framework for the empirical analysis. In Section 4.4 we outline the data-construction process and provide relevant summary statistics. In Section 3.5, we present our empirical model and estimation strategy. Section 3.6 gives a description and discussion of the empirical findings, along with the policy implications thereof. The final Section 3.7 summarizes and concludes the limitations of the study, and potential directions for further research.

3.3 Conceptual framework

3.3.1 A parametric stochastic frontier approach to eco-efficiency

In order to analyze the economic-ecological performance of businesses, we build upon and further develop the frontier setting developed by Kuosmanen & Kortelainen (2005) and Orea & Wall (2017). We base our approach on the definition of eco-efficiency from the literature on ecological economics. Eco-efficiency is defined as the ratio between economic performance y (traditionally *value added*) and environmental damage (D). Suppose we observe a set of $k = 1, \ldots K$ comparable production units which generate economic output y each year in period $t = 1, \ldots T$ through $n = 1, \ldots N$ environmental pressures $\mathbf{s^t} = (s_1^t, \ldots, s_N^T)$ that damage the ecological system.

There are numerous ways to generate economic output which vary in their environmental impact (Kuosmanen & Kortelainen, 2005). Therefore, we introduce the time-dependent pressure-generating technology set (Kortelainen, 2008; Kuosmanen & Kortelainen, 2005; Picazo-Tadeo et al., 2012):

$$PGTS_t = \left[(y_t, \mathbf{s}_t) \in \mathbb{R}_+^{N+1} \middle| \text{ economic output } y_t \text{ can be generated with ecological pressures } \mathbf{s}_t \right] \tag{3.1}$$

It contains all technically and economically feasible combinations of economic outcome (y) and environmental pressures ($\mathbf{s}$) in period t. It is further assumed that all production units have a common underlying pressure-generating technology.

Note that this concept relates to economic and environmental pressures in a non-physical sense. Greenhouse gas emissions (and other environmental pressures) are therefore looked at from an impact-based point of view as opposed to a quantity-based point of view. As Kuosmanen & Kortelainen (2005) state, this is a fundamental presupposition for the eco-efficiency concept. Physical inputs and outputs that affect economic outcome and those that have an impact on the environment are implicitly included in y and $\mathbf{s}$. This is different from more traditional environmental economics approaches that use inputs and outputs as key elements of their models.[1]

Eco-performance

Most studies (e.g. Kuosmanen & Kortelainen, 2005; Pérez Urdiales et al., 2016) base their analyses on the eco-efficiency concept, which is defined as the ratio of economic value added and environmental pressures. Leaning on this approach, we define eco-performance (EP) at time t as follows:

$$EP^t = \frac{\text{Economic output}^t}{\text{Environmental pressure}^t} = \frac{y^t}{D(s_1^t, s_2^t, \ldots, s_N^t)} \tag{3.2}$$

where $D(s_1^t, s_2^t, \ldots, s_N^t)$ is a function that reflects the environmental damage associated with the individual ecological pressures/environmental stress factors **s** at time t.

As can be seen from (3.2), we depart from the traditional definition of eco-efficiency in several ways. In line with Kortelainen (2008) and Picazo-Tadeo et al. (2014), we define the above-mentioned ratio as *eco-performance*. Eq. 3.2 does not deliver any benchmark or baseline for which to compare given levels of performance as assumed by the relative concept of eco-efficiency. Therefore to clearly distinguish between the absolute concept presented in (3.2) and eco-efficiency, we choose the term eco-performance (Kortelainen, 2008).[2] What is more, we define eco-performance in a more general sense, in that we do not restrict the economic outcome variable to be value added. We select farm revenues (r) as the economic outcome variable of choice. This is mainly because we argue that revenues are

[1] In that sense, the presented concept deviates from multiple analyses on production in the presence of undesirable outputs or by-products, respectively (e.g. Atkinson & Tsionas, 2018; Førsund, 2018; Malikov et al., 2018). In these studies, pollutants such as nitrogen surpluses or carbon dioxide (CO_2) emissions are considered physical quantities. Put differently, the approach presented in this study rather relates to the field of ecological economics as opposed to environmental economics, where emission quantities are assumed. Here, we seek to relate ecological pressures to economic outcome, which is distinctly different from the above-mentioned approaches that seek to model the relationship between physical inputs and outputs of a production process. In our approach, we do not explicitly model the input-output relationship. Note however, through relating ecological pressures to economic outcome the underlying production technology is implicitly contained in the model, although not explicitly modeled. For instance, GHG emissions are a result of the use of inputs such as fertilizers or livestock. Hence, these inputs are reflected by GHG emissions in our PGTS.

[2] Throughout this study, we use the terms eco-performance, environmental performance as suggested by Kortelainen (2008) and eco-productivity interchangeably. We defer from using the term 'environmental productivity' to avoid confusion, as this term has previously been used in different contexts (e.g. Ball et al., 2004; Managi, 2006; Managi et al., 2005)

economically more relevant to farmers than economic value added. Given binding regulatory and expenditure constraints with respect to inputs, Chambers & Lee (1986), and Kumbhakar & Bokusheva (2009) show that most farms aim at maximizing revenues (or equivalently outputs).

As highlighted by Kortelainen (2008), the absolute value of eco-performance is not very informative and hardly interpretable as such. Therefore, we will focus on the assessment of change rates in eco-performance by making use of an eco-performance index. We will return to this issue in Section 3.3.2.

The stochastic frontier pressure conversion model

Representing the pressure-generating technology (3.1) by a functional form, we obtain a function that provides the relationship between environmental stress **s** and economic outcome y, where, for now, no inefficiency is assumed:[3]

$$y = D(s_1,\ s_2,, \ldots, s_N; \boldsymbol{\beta}) \tag{3.3}$$

aside from the known elements in the equation, $\boldsymbol{\beta}$ represents a parameter vector to be estimated. We call (3.3) a *pressure conversion function* as it describes how ecological pressures (independent variables) are converted to economic outcome (dependent variable). Hence, economic outcome is expressed as a function of ecological damage caused by environmental pressures. As Orea & Wall (2017) note, the parametric specification of Eq. 3.3 allows us to evaluate the marginal contribution of pressure s_n to the economic outcome y of production unit k at time t. For instance, it is possible to assess by how much the economic output decreases if the n-th pressure is decreased by one unit (percent), ceteris paribus. Thus, we can also assess which environmental pressure may be most costly to reduce for a given firm k. A major concern could be that the omission of capital and labor in our framework might lead to ommited variable bias (OVB) since capital and labor are commonly modeled as key inputs to production technologies. If labor and capital are fixed over the analyzed time period, the specified fixed-effects model can reliably remove these variables. In farming, labor and capital are often by definition fixed in the long-term. This assumption appears to be rather sensible in the Bavarian farming context. What is more, OVB occurs only

[3]To avoid notational clutter, we omit the subscripts for time $t = 1, \ldots T$ and the production units $k = 1, \ldots K$.

when the omitted variables are both correlated with the independent variables as well as with the dependent model. We do not expect this to be a major problem in our application. We included a robustness check in the Appendix 3.8.6, where we estimated alternative specifications of the *pressure conversion function* that included capital and labor as independent variables. As expected, we could not find indication for the presence of omitted variable bias due to these two variables.

Eq. 3.3 neglects the fact that not all farms generate the economically feasible maximum level of revenue, given their level of environmental damage. To account for (ecological) inefficiency and statistical noise, (3.3) can be expressed as a stochastic frontier following Aigner et al. (1977) and Meeusen & van den Broeck (1977):

$$y = D(s_1,\ s_2, , \ldots, s_N; \boldsymbol{\beta})\ e^{-u+\nu} \tag{3.4}$$

where $\epsilon = -u + \nu$ represents a composite error term consisting of an ecological inefficiency component (u) and statistical noise (ν). We assume that environmental damage is monotone in the sense that it cannot decrease, ceteris paribus, if any environmental pressure is increased. Since $y = D(\boldsymbol{s})$, monotonicity must also hold for Expression 3.3, which is a necessary condition for obtaining a sensible interpretation of individual (eco-)efficiency scores (Henningsen & Henning, 2009; Sauer et al., 2006). One major advantage of choosing a parametric approach over DEA is the fact that we can account for stochastic noise and thus take into account the effect of random shocks, outliers and measurement errors. Contrary to previous DEA-based studies on eco-efficiency, our parametric approach allows us to specify $D(\boldsymbol{s})$ in a nonlinear fashion. This relates to the relationship between environmental pressures and ultimate environmental damage. Kortelainen (2008, p.703) states "[. . .] *the relationship between the environmental pressure and the ultimate environmental impact can be complex, nonlinear, and very difficult to predict"*. In that sense, the presented framework might better grasp underlying nonlinearities.

Returns to pressure scale

There are two major perspectives concerning the question if the pressure-generating technology exhibits constant or variable returns to scale (VRS) (Picazo-Tadeo et al., 2012). From an ecological point of view, farming could be considered a constant returns to scale (CRS) activity as the effect of farming on the environment is directly related to agricultural practices rather than to the allocation of land and other inputs into individual farms (Picazo-Tadeo et al., 2012, 2011). However, from an economic point of view, returns to pressure scale of farming activities could also be considered variable because ecological pressures often depend on the use of inputs both in terms of quantity as well as intensity. As the input-output relationship in farming is usually described by VRS, this is also likely for returns to pressure scale (Bonfiglio et al., 2017). As the imposition of CRS would *a priori* imply linearity between emissions and economic outcome, we decide to relax the more restrictive assumption of constant returns to pressure scale. In allowing returns to pressure scale to be variable, we can actually empirically test which of the two perspectives on pressure scale is more realistic.

Eco-efficiency

In the spirit of Shepard's technical efficiency (TE) concept, we define eco-efficiency as the ratio between the observed economic output and the frontier economic output:

$$EE = \frac{y}{D(\boldsymbol{s}'\boldsymbol{\beta})\ e^{\nu}} = \frac{D(\boldsymbol{s}'\boldsymbol{\beta})\ e^{-u}\ e^{\nu}}{D(\boldsymbol{s}'\boldsymbol{\beta})\ e^{\nu}} = e^{-u} \tag{3.5}$$

This economic outcome-oriented measure of eco-efficiency by definition lies between zero and one.[4] It measures the economic output of the k-th firm relative to the maximum attainable economic outcome by a fully-efficient firm for a given level of environmental damage. Note that this specification of eco-efficiency is different from the traditional one given at the beginning of this section, which is defined as the ratio between economic output and aggregate environmental pressure.

3.3.2 Eco-performance – dynamics and decomposition

As earlier mentioned, in order to obtain a more comprehensive understanding of farms' ecological performance, we introduce the concept of eco-performance change which builds upon the literature on total factor productivity (TFP) (mainly Balk, 2001; Coelli & Rao, 2005b; Kumbhakar & Lovell, 2000; Orea, 2002). Analogous to Orea (2002), our eco-performance analysis is largely based on the Malmquist

[4] We choose an outcome-oriented over a emission-oriented measure because we argue that in most cases it is more realistic to assume that farms seek to maximize economic output for a given level of environmental damage than to minimize emissions for a given level of output.

TFP index and its three differentials: scale change, technical change and efficiency change. Our approach deviates from Orea (2002) in the sense that we do not consider TFP but rather total eco-pressure productivity (or in short: eco-performance).

Caves et al. (1982) measure TFP change based on Malmquist input or output distance functions, respectively. Basically, the Malmquist TFP change index is defined by the ratio of the distances of each data point relative to a common technology. Here, the common technology is respresented by the pressure conversion frontier (3.4). Furthermore, we make use of an output-oriented index, i.e. the vector of ecological pressures (s_t) is assumed to be fixed while economic outcome (y) is maximized given the pressure-generation technology. Eco-performance change between two periods f and t can then be expressed as the geometric mean of the Malmquist indices for periods f and t. Furthermore, we assume that some degree of eco-inefficiency can be observed for most production units, i.e. $d_0^f(y_f, s_f) \leq 1$ and $d_0^t(y_t, s_t) \leq 1$. Hence, the conversion of environmental pressures to economic output are additionally subject to efficiency changes (Coelli & Rao, 2005b). We can express eco-performance change as:

$$m_0(y_f, s_f, y_t, s_t) = \frac{d_0^t(y_t, s_t)}{d_0^f(y_f, s_f)} \times \left[\frac{d_0^f(y_t, s_t)}{d_0^f(y_f, s_f)} \times \frac{d_0^t(y_t, s_t)}{d_0^t(y_f, s_f)} \right]^{1/2} \tag{3.6}$$

where $d_0^f(y_t, s_t)$ stands for the distance from the period t observation to the technology in f and the second term of the Eq. 3.6 reflects changes in the underlying pressure generating technology over time, while the term outside the square brackets reflects eco-efficiency change. The presented Malmquist index does not make any assumptions with respect to returns to scale. However, as mentioned by multiple authors, Eq. 3.6 is only true if the technology exhibits constant returns to scale. Among others, Balk (2001) and Orea (2002) state that productivity can also be improved through improvements in the scale of operations, a component that the Malmquist index defined by Caves et al. (1982) does not capture. As noted earlier, we explicitly allow for VRS. It is therefore pivotal to add a scale component to the eco-performance index. Consequently, we make use of the widely accepted and utilized generalized Malmquist TFP index suggested by Orea (2002), which takes account of scale economies and complies with the requirements of identity, separability and monotonicity:[5]

$$\ln g_0 = \ln m_o + \frac{1}{2}\sum_{n=1}^{N}\Bigg[\Bigg(-\sum_{n=1}^{N}\frac{\partial \ln D_0(t)}{\partial \ln s_n} - 1\Bigg) \times e_n(t) + \Bigg(-\sum_{n=1}^{N}\frac{\partial \ln D_0(f)}{\partial \ln s_n} - 1\Bigg) \times e_n(f)\Bigg] \times \ln\Bigg(\frac{s_n^t}{s_n^f}\Bigg) \tag{3.7}$$

with

$$\text{Returns to pressure scale}^t = e_n(t) = \frac{\partial \ln D_0(t)/\partial \ln s_n}{\sum_{n=1}^{N} \partial \ln D_0(t)/\partial \ln s_n}$$

Further information regarding the operationalization of the productivity change decomposition by Orea (2002) is given in Section 3.5.

3.3.3 GHG emissions and eco-performance

In the context of this study, we wish to focus specifically on GHG emissions. For that purpose, we define ecological damage in terms of GHG emissions as a function of the three major pressures on the climate, namely carbon dioxide (s_{CO_2}), methane (s_{CH_4}) and nitrous oxide (s_{N_2O}). Contrary to that, Kuosmanen & Kortelainen (2005) argue in favor of using only one aggregate measure for greenhouse gases as they contribute the same environmental problem, namely the greenhouse effect. We, however, believe it is important to decompose climatic stress into the above-mentioned pressures for two reasons. First, despite the fact that GHGs contribute to the same effect, they do have diverging relative impacts on the climate and thus on the ecological damage they cause, which should be accounted for. Second, as stated earlier, physical input and output quantities are implicitly included in D. In the farming context, they are linked to emissions through different activities and management practices. For instance, nitrous oxide (N_2O) emissions are associated with the application of nitrogen fertilizers, while methane (CH_4) primarily stems from the digestive system of ruminants. If we seek to evaluate these pressures in terms of the climate stress they produce, we implicitly evaluate

[5]In general, there is no consensus in the literature on what productivity change measure to use best. While some TFP indices such as the Fisher or Tornquist index require price information, others such as the Malmquist or Hicks-Morsteen index do not depend on prices. O'Donnell (2012) argues in favor of the Hicks-Morsteen-Index. Balk (2001) and Orea (2002) advocate the use of the Malmquist index, which has experienced quite extensive use in empirical work (e.g Coelli & Rao, 2005a; Frick & Sauer, 2018; Song et al., 2016). What is more, Briec & Kerstens (2011) show in their study that there are only minor numerical discrepancies between the Malmquist and Hicks-Morsteen-Index.

different farming activities, a fact we would miss if we used an aggregated measure of GHGs.

Analogous to the *pressure conversion frontier*, by inserting GHG emissions into Eq. 3.4 and using firm revenue (r) as economic outcome variable, we obtain the *'GHG conversion frontier'*:

$$r = D(s_{CO_2},\ s_{CH_4},\ s_{N_2O}, t; \boldsymbol{\beta})\ e^{-u+\nu} \tag{3.8}$$

Based on (3.8), all metrics and concepts presented in Sections 3.3.1 and 3.3.2 are equally true for GHG emissions. This allows us to assess firms' returns to emission scale (RTES), emission efficiency as well as eco-performance with respect to GHG emissions. Variable t reflects the time to properly capture the time-dependency structure of marginal emission damage (Field & Field, 2009).

It is largely undisputed that anthropogenic global warming through the emission of greenhouse gases has detrimental environmental impacts. However, according to Weitzman (2012) it is barely possible (if not impossible) to properly quantify high-temperature damage. GHG emissions are not environmentally harmful per se if they do not exceed the carrying capacity of the atmosphere. However, starting from that threshold, additional GHG emissions can lead to erratic climate conditions and, by extension, to ecological harm. The exact relationship between greenhouse gas releases and high-temperature damage $D(s_{CO_2}, s_{CH_4}, s_{N_2O})$ is fairly unknown and is seen as highly complex and nonlinear (Kuosmanen & Kortelainen, 2005). This is why we cannot attribute the adverse effects of climate change to a specific farm in practice. As such, and in line with Kuosmanen & Kortelainen (2005), we do not seek to compute the ultimate environmental impacts of farms' ambient GHG releases through climate change, and instead stick to the level of environmental pressures or climate stress factors, respectively.

3.4 Data

As will be seen in the following, data requirements for estimating the above-mentioned conceptual model largely deviate from most other studies in the field of production economics. First, the model has no direct recourse to physical inputs and outputs. Second, besides farm accountancy data, we need several additional datasets for our analysis. There is no database of farm-level greenhouse gas emissions in Europe, nor can GHG emissions easily be retrieved from farm accountancy data only, although such data serve as an important source for calculating farm-level GHG emissions.

3.4.1 Farm accountancy dataset

The most important dataset of our empirical analysis is farm accountancy data for Germany, and more specifically for the federal state of Bavaria. This dataset is part of the European Farm Accountancy Data Network (FADN). Data are annually collected from approximately 3,100 farms. It is an unbalanced farm-level panel dataset and participation is voluntary. The sample is stratified with respect to farm location, size classes, and specialization of the farms. In addition to financial records, socio-economic information is provided such as the education level of the farm manager, number of household members or the on-farm labor structure. The sample covers the time period from 2005 through 2014.

Although our empirical example is based on a regional sample of farms, we believe that our empirical findings are relevant in a larger European context. The case of Bavaria may be regarded as somewhat representative for other European regions as well. For instance, farms in Bavaria managed on average 34.7 ha of land in 2014 which is similar to average farm sizes in, e.g., Ireland, Belgium and the Netherlands (European Statistical Office, 2020). Also, Bavarian dairy farmers kept on average 34.1 livestock units (LUs) of dairy cows between 2005 and 2014. This value lies only slightly above the average for all European regions (30.4). On average, Bavarian pig farms managed 154.4 LUs of pigs, which was approximately equal to the European average of 154.8 LUs. As for the labor structure, on average 1.6 average work units (AWUs) worked on Bavarian farms between 2005 and 2014, while the European average was 1.5 AWUs in the same period (European Statistical Office, 2020).

For this research, we focus on four important farm types in European agriculture, namely dairy farms, pig fattening farms, mixed farms (livestock and crop) and crop farms. Farms are categorized according to their principal activity within the farm business. Specialized farms (dairy, pig and crop) are assigned to the respective farm type if the output share of their characteristic produces exceed 66% in total revenues (milk, fattening pigs, grains). As for mixed farms (i.e. crop-livestock systems), no primary product accounts for more than 66 % of total revenues. Only farms are included that provide information for at least 3 years. Organic farms were excluded from the sample based on the assumption of a fundamental different technology compared to conventional farms. In our application, farm revenue enters the model as the economic outcome variable, which is defined as the value of sales (taxes included, subsidies excluded). In order to eliminate price effects from our analysis we deflate revenues from different farming activities to the base year 2014. Price indices of agricultural producer prices provided by the German Statistical Office are used for deflating the data. Table 3.1 gives an overview of the subsamples used for the empirical analysis and the respective deflated revenues.

Table 3.1: Subsample description - farm types (2005 - 2014)

	Dairy	**Swine**	**Mixed**	**Crop**
Number of observations	9,574	3,796	2,558	5,318
Number of farms	1,513	585	711	919
Avg. participation (years)	6.33	6.49	3.6	5.79
Revenue (Const. 2014 Euro)				
- Mean	154,453	296,348	172,499	109,261
- SD	88,074	205,914	143,274	115,343
- Min	11,087	700	403	241
- Max	1,139,323	1,534,219	1,464,453	1,442,213

3.4.2 Construction of farm-level GHG emissions and additional data

As agricultural GHG emissions are considered a nonpoint source pollution, direct measurement is rather impractical and costly when applied to a large number of farms (Dick et al., 2008; Paustian et al., 2004; Smith et al., 2008). Therefore researchers have frequently made use of indirect methods to estimate agricultural GHG emissions, where emissions are regarded as the outcome of a combination of farm activities and management practices (Baldoni et al., 2018). In a multitude of studies, greenhouse gas emissions are approximated by making use of the ISO-standardized life cycle assessment inventory approach, where emissions associated with each stage of the production change are cumulated (ISO 14044:2006, 2006). These studies, however, only consider small samples of specific farm types and rarely consider their evolution over time. This is largely due to the expensive data-collection procedure.

An alternative procedure for recovering farm-level GHG emissions is based on the guidelines of the United Nations Framework Convention on Climate Change

(UNFCCC) and the Intergovernmental Panel on Climate Change (IPCC), which are designed for national GHG inventories, and are thus considered an internationally accepted and widely applicable standard (e.g. Casey & Holden, 2005; Olesen et al., 2006). Furthermore, we use the methodological extensions to the IPCC method added by Haenel et al. (2018). A multitude of studies has shown that the approach is also well-suited for retrieving GHG emissions at the micro-level (Baldoni et al., 2017, 2018; Coderoni & Esposti, 2018; Dick et al., 2008). While earlier studies were still relatively limited with respect to time and sample scope, recent advances towards integrating farm accountancy data allow larger samples and comparisons across farms and time.

The basic idea is that GHG emissions at farm level can be retrieved from the following relationship:

$$Em_{it} = \sum_{l=1}^{L}(AD_{itl} \times EF_{tl}) \tag{3.9}$$

where Em_{it} describes total GHG emissions of farm k at time t. AD_{itl} refers to some activity data of farm k at time t, referring to emission source $l = 1, \ldots, L$. AD_{itl} describes the AD (e.g. number of animals, fertilizer quantities or diesel consumption) of the l-th activity at time t, while EF_{tl} stands for (implied) emission factors (EFs) which characterize the emission quantities associated with a specific activity (e.g. kg CH_4 released per dairy cow per year, or kg N_2O associated with application of 1 kg of mineral fertilizer).

We define relevant system boundaries at the farm-gate to keep the inventory transparent and comprehensible. Also, procedures within the farm-gate leading to emissions are directly within the control of the farms and can therefore unambiguously be attributed to farmers' performance (Coderoni & Esposti, 2018). In terms of which emission sources should be included in the inventory, first and foremost, we make use of the Common Reporting Format Sector 3 "Agriculture" of the United Nations Framework Convention on Climate Change (UNFCCC, 2014). On top of that, CO_2 emissions from the use of energy, more specifically from on-farm fuel combustion is included, which refers to the UNFCCC "Energy" sector. An overview of the major emission sources, AD and EFs is given in Table 3.2. We distinguish between the three major categories: livestock, crop cultivation and energy use. As for livestock, animal numbers are retrieved from the farm accountancy

Table 3.2: Summary of greenhouse gas sources, activity data and utilized emission factors for the computation of farm-level GHG emissions.

Gas	Emission source	Activity data		EF
		FAD	Other	
	Livestock			
CH_4	Enteric fermentation	Livestock count		regional
CH_4	Manure management	Livestock count		regional
N_2O	Manure management (direct and indirect)	Livestock count	NH_3 and NO emission factors (indirect)	regional
	Crop cultivation			
N_2O	Use of synthetic fertilizers	Fertilizer expenditures	State-level shares, prices	regional
N_2O	Use of organic fertilizers	Livestock count	N excretion factors	regional
N_2O	Atmospheric deposition of reactive nitrogen	Livestock count & fertilizer expenditures	$Frac_{Gas}$	default & regional
N_2O	Leaching and surface runoff	Livestock count & fertilizer expenditures	$Frac_{Leach}$	default & regional
N_2O	Crop residues	Crop area and yield	Various constants	default
CO_2	Urea application	Fertilizer expenditures	State-level shares, prices	regional
CO_2	Liming	Fertilizer expenditures	State-level shares, prices	regional
	Energy use			
CO_2	Fuel combustion	Fuel expenditures	Diesel price	default
N_2O	Fuel combustion	Fuel expenditures	Diesel price	default
CH_4	Fuel combustion	Fuel expenditures	Diesel price	default

dataset. For a more precise calculation, animal categories such as cattle, pig and poultry were further decomposed according to Haenel et al. (2018).[6]

While GHG emissions from livestock keeping can fairly easily be recovered from farm accountancy information, this is not true for crop cultivation, as additional information is needed for retrieving AD. The major challenge lies in the fact that

[6]The following distinction has been made with respect to nitrogen excretion rates as well as enteric fermentation rates: (a) Horses: heavy horses, ponies and light horses (b) Cattle: calves, male beef cattle, heifers, mature males > 2 years, dairy cows and suckler cows (c) Pigs: weaners, fattening pigs, boars, sows (d) Sheep (e) Poultry: laying hens, pullets, broilers, other poultry.

the prime purpose of farm accountancy data is to record production, and rarely the evaluation of ecological performance. Figure 3.5 in the appendix illustrates how additional data sources can be utilized beside the farm accountancy data to calculate GHG emissions.

As already found by Coderoni & Esposti (2018), a major challenge is to recover physical fertilizer quantities when the dataset only provides information on fertilizer expenditures. Coderoni & Esposti (2018) use a fixed conversion factor for fertilizer expenditures to retrieve quantities. We argue against this approach, since quantities of, say, nitrogen vary in terms of their price and their relative share within all fertilizers used over time. Therefore, as a means to recover the individual quantities of synthetic N application, we construct a double-weighted mean in a 4-step procedure.[7] A description of the method as well as more details on the emission sources can be found in the appendix of this paper. The interested reader is referred to Haenel et al. (2018), who describe and discuss the IPCC method in detail for the German context. There, all underlying computation formulas for retrieving the respective AD can be found.

3.4.3 Post-processing, GHG emissions and emission intensities

Although the IPCC method for reconstructing GHG emissions as described in Section 3.4.2 is widely accepted in the literature, the fact that farm-level GHG emissions are computed by utilizing data from different aggregation levels inevitably leads to some anomalies in the data. For this reason, we apply the blocked adaptive computationally efficient outlier nominators (BACON) algorithm as suggested by Billor et al. (2000) to detect multivariate outliers. In total, 580 observations are excluded (162 dairy, 109 pigs, 83 mixed, 179 crop) from the analysis, i.e. 2.51 per cent of all observations across the four farm types.[8]

Summary statistics of the GHG releases per farm can be found in Table 3.3. To ensure comparability, CH_4 and N_2O emissions were converted to CO_2-equivalents (CO_{2eq}). To that end, N_2O and CH_4 quantities were multiplied by their respective global warming potentials (298 and 34, respectively) as per the IPCC's Fifth Assessment Report (IPCC, 2013).[9] There are clear differences across the inspected

[7] This also applies to urea and calcium carbonate (CaCO). We obtained diesel quantities in a similar fashion, which allows us to recover CO_2 emissions from fuel combustion.

[8] For crop farms, we did not consider CH_4 in the outlier detection procedure, since 75% of that subsample did not produce any CH_4 at all in the period 2005-2014.

[9] Considering the inclusion of climate carbon feedback and a 100-year time horizon.

farm types. On average, dairy farms have the highest GHG emission rate per year (417.1 t CO_{2eq}) followed by mixed and pig farms. Crop farms show the smallest number of total GHG emissions per farm. As for dairy farms, unsurprisingly, methane is the major contributor to total emissions. This is also true for mixed farms. For the other types, it is N_2O. CO_2 is present in all farming systems but plays a lesser role with respect to total emissions apart from crop farming. It is worthwhile to note the large standard deviations concerning all farm types and greenhouse gases. This alludes to a high level of heterogeneity across individual farms. These first descriptive results are comparable to other findings, e.g. by Baldoni et al. (2017) and Coderoni & Esposti (2018) for Italian farms.

To get a first impression of farms' capacity to convert GHG emissions to monetary output, we compute emission intensities (EI, compare Baldoni et al., 2018). EI measures emissions in kg CO_{2eq} per unit of output expressed in constant Euros. Table 3.3 shows that dairy farms have the highest EI of 3.81 kg $CO_{2eq}/€$, i.e. on average a Bavarian dairy farm emits 3.81 kg CO_{2eq} to produce one Euro of output. Mixed farms are next (2.03 $CO_{2eq}/€$), followed by crop farms (1.47 $CO_{2eq}/€$) and finally pig fattening farms(1.05 $CO_{2eq}/€$).

Table 3.3: Farm-level GHG emissions and emission intensities of Bavarian farms.

	Dairy		**Swine**		**Mixed**		**Crop**	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Methane (Tonnes CO_{2e})	417.1	204.8	108.6	78.8	154.1	160.4	8.1	38.5
Nitrous Oxide (Tonnes CO_{2e})	88.4	48.6	116.2	70	97	67.4	68.1	67.9
Carbon Dioxide (Tonnes)	29	20	37.8	26.9	44.9	31.9	36.7	33.3
Total (Tonnes CO_{2e})	534.5	267.2	262.6	160.7	296	234.3	112.9	113.5
Emission Intensity (kg $CO_{2e}/€$)	3.81	0.79	1.05	0.42	2.03	1.14	1.47	3.16

3.5 Empirical specification

3.5.1 Estimation strategy

In our study we employ the stochastic frontier approach as first proposed by Aigner et al. (1977) and Meeusen & van den Broeck (1977). The production economics literature delivers a wide range of candidate stochastic frontier models for the estimation of the GHG conversion frontier (3.8) in a panel data setting; for a recent overview see e.g. Kellermann (2015) and Kumbhakar et al. (2014). Examples for regularly used stochastic frontier models are the 'error components model' by Battese & Coelli (1992), the 'efficiency effects' approach by Battese & Coelli (1995) and the 'true fixed/random effects' model by Greene (2005). We decide to use the stochastic frontier specification introduced by Kumbhakar et al. (2014, KLH) and write (3.8) as:[10]

$$r_{knt} = \alpha_0 + D(\boldsymbol{s}_{knt}, \boldsymbol{\beta}) + \mu_k + \nu_{kt} + \eta_k - u_{kt} \tag{3.10}$$

Here, the error consists of four components. $\eta_k > 0$ and $u_{kt} > 0$ account for inefficiency, whilst μ_k and ν_{kt} represent firm effects and noise, respectively. As Kumbhakar et al. (2014) note, all of these components have already been implemented in one way or another, yet not all at the same time. Contrary to most commonly utilized stochastic frontier models (e.g. Chen et al., 2014; Greene, 2005; Wang & Ho, 2010), the KLH model properly distinguishes between time-varying (u_{kt}) and time-invariant inefficiency (η_k), while also accounting for the effect of unobserved farm heterogeneity on outcome.[11]

Kumbhakar et al. (2014) suggest a multi-step estimation procedure for the KLH model. To apply this, (3.10) must be operationalized by rewriting it as follows:

$$r_{kt} = \alpha_0^* + D(\boldsymbol{s}_{kt}, \boldsymbol{\beta}) + \alpha_k + \epsilon_{kt} \tag{3.11}$$

[10] Kumbhakar et al. (2014) mention various shortcomings of earlier stochastic frontier models. For instance, Battese & Coelli (1992) only allow inefficiency to change over time exponentially. Also, firm effects (fixed or random) are often not clearly distinguished from (persistent) inefficiency. The KLH model overcomes many of the problems associated with previous stochastic frontier models as shown below.

[11] Time-varying (transient) inefficiency is related to time-varying issues such as the adaptation to changes in the firms' environment and is therefore a short-term concept which changes over time. It is related to operative business activities. Persistent inefficiency is a long run concept and stable over time. It is related to the structure of a company. If a company is persistent (structurally) inefficient this may require the introduction of several policy measures such as change in ownership, increasing/reducing the size of a company etc.

where $\alpha_0^* = \alpha_0 - E(\eta_k) - E(u_{kt})$; $\alpha_k = \mu - \eta_k + E(\eta_k)$ and $\epsilon_{kt} = \nu_{kt} - u_{it} + E(u_{kt})$. Here, α_k and ϵ_{kt} have zero mean and constant variance. Technical details regarding the estimation of Eq. 3.11 can be found in the appendix and in Kumbhakar et al. (2015).

As we wish to focus on GHG emissions in this study, methane ($s_{CH_4,kt}$), nitrous oxide ($s_{N_2O,kt}$) and CO_2 ($s_{CO_20,kt}$) enter in the operationalized GHG conversion function (3.11) as independent variables. Revenue r_{kt} is the dependent variable. In terms of the functional form of the *'pressure/ghg conversion function'* (3.8), Orea & Wall (2017) suggest a weighted mean of the environmental pressures (**s**) to depict ecological damage caused by these pressures. We argue against this approach in that it appears too simplistic to properly describe the unknown true relationship between environmental pressures (here: emissions) and ecological damage. As stated earlier, the relationship between emissions and damage is assumed to be highly complex and nonlinear (Kuosmanen & Kortelainen, 2005). To adequately account for this nonlinearity and complexity and to constrain our functional form as little as possible, we suggest the use of the second-order flexible translog functional form for $D(\boldsymbol{s})$. We add a time variable to properly account for the panel structure of the data and to accommodate the time-specific character of the climate damage function as proposed by Field & Field (2009). Hence, we specify the empirical stochastic GHG conversion frontier (3.10) as:

$$\begin{aligned}\ln r_{kt} =& \alpha_0 + \sum_{n=1}^{N} \alpha_n \ln s_{nkt} + \frac{1}{2}\sum_{n=1}^{N}\sum_{l=1}^{N} \alpha_{nl} \ln\ s_{nkt} \ln s_{lkt} \\ &+ \beta_t t + \frac{1}{2}\beta_{tt}\ t^2 + \sum_{n=1}^{N} \beta_{nt}\ t\ \ln s_{nkt} + \mu_k + \nu_{kt} + \eta_k - u_{kt}\end{aligned} \tag{3.12}$$

where $\alpha_0, \alpha_n, \alpha_{nl}, \beta_t, \beta_{tt}$ and β_{nt} are unknown parameters to be estimated. t represents time.[12] The multi-step procedure for estimating Eq. 3.12 is conducted separately for each farm typology, since it is assumed that they relate to distinct reference technologies. Partial outcome elasticities of the greenhouse gases can be retrieved in the standard fashion:

$$e_{nkt} = \frac{\ln r_{kt}}{\ln s_{nkt}} = \alpha_n + \sum_{l=1} \alpha_{nl}\ \ln s_{lkt} \tag{3.13}$$

[12]We provide an alternative specification including capital and labor as a robustness check (see Appendix. 3.8.6).

Returns to emission scale (RTES) are the sum of the partial elasticities:

$$\mathrm{RTES}_{kt} = \sum_{n=1} e_{nkt} \; \ln s_{nkt} \tag{3.14}$$

In order to evaluate the validity of our models, we test several alternative specifications against our baseline model (3.12). First, we test the hypothesis that there is in fact an ecological-technical change with respect to emissions over time. Second, we test if technical change might be better described as Hicks-neutral. Third, we test the translog formulation against the Cobb-Douglas functional form.

3.5.2 Computation of productivity dynamics

As described in Section 3.3.2 and in line with Orea (2002), the eco-performance change (EPC) is composed of three components, emission efficiency change (EEC), ecological-technical change with respect to emissions (ETC) and emission scale change (ESC). The EEC index can be computed as follows (Coelli & Rao, 2005b):

$$EEC_{k(ft)} = \frac{EE_{kt}}{EE_{kf}} \tag{3.15}$$

ETC between two adjacent periods can be calculated as the geometric mean of the partial derivatives of (3.12) with respect to time:

$$ETC_{i(ft)} = \exp\left(\frac{1}{2}\left[\frac{\partial \ln r_{if}}{\partial f} + \frac{\partial \ln r_{it}}{\partial t}\right]\right) \tag{3.16}$$

Finally, in accordance with Orea (2002), emission scale change is computed as:

$$ESC_{i(ft)} = \exp\left(\frac{1}{2}\sum_{n=1}^{N}\left[e_{nif}\mathrm{SF}_{if} + e_{nit}\mathrm{SF}_{it}\right]\ln\left(\frac{s_{nit}}{s_{nif}}\right)\right) \tag{3.17}$$

where $SF_{if} = (e_{if} - 1)\,/e_{if}$; $e_{if} = \sum_{n=1}^{N} e_{nif}$ and $e_{nif} = \partial \ln r_{is}/\partial \ln s_{if}$. Total emission-performance change (or emission productivity change) is the sum of the single components:

$$\mathrm{EPC}_{i(ft)} = \mathrm{EEC}_{i(ft)} + \mathrm{ETC}_{i(ft)} + \mathrm{ESC}_{i(ft)} \tag{3.18}$$

3.6 Empirical results and discussion

3.6.1 Model specification

Parameter estimates of the empirical greenhouse gas conversion function (3.12) were obtained by estimating a fixed effects regression model for each farm type separately. Hausman tests were conducted, which rejected both the random effects model and the pooled regression model at the 1% significance level for all farm types. Cobb-Douglas specifications were rejected in favor of the translog model. Furthermore, Wald tests rejected the hypotheses of zero technical change as well as Hicks-neutral technical change. Therefore we opted for specifications that allow for non-monotonic and non-neutral technical change. Estimation results of the favored models are presented in Table 3.4.

All first-order coefficients can be interpreted as partial elasticities of GHGs at the sample mean, which are positive and significantly different from zero for all farm types and greenhouse gases. For instance, a 1% decrease in CH_4 is, ceteris paribus, associated with a 0.68% decrease in revenues for an average dairy farm. Evaluating at the sample mean, we can find that the same GHGs have a different association with revenues depending on the farm type. Reducing CH_4 is most costly for dairy farms, in that their revenues decreased most significantly, compared with N_2O and CO_2. The same applies to pig fattening farms; however less pronounced. For mixed and crop farms, it is most expensive to reduce CO_2 followed by nitrous oxide, while methane only plays a minor role.

Steps two and three were estimated based on the results provided in Table 3.4. Maximum likelihood estimations were conducted according to (3.19) and (3.20). In order to avoid biased standard errors, we computed bootstrapped standard errors. This allows us to evaluate the statistical significance of the estimated parameters. Estimation results can be found in Table 3.5 and Table 3.6. As can be seen, all relevant coefficients have small standard errors and are statistically significant, indicating that both permanent and time-varying efficiencies are present in all farm typologies. Additionally, likelihood ratio tests were conducted to test the stochastic frontier models against ordinary least squares (OLS), where OLS were rejected in favor of the frontier models for all farm types.

As shown by Sauer et al. (2006) and Henningsen & Henning (2009), the monotonicity condition plays an important conceptual role in stochastic frontier analyses. In order to sensibly interpret (partial) elasticities, efficiency scores and other derived

Table 3.4: KLM step 1 result table – fixed effect regression for 4 different farm types in Bavarian agriculture.

	Dependent variable:			
	$\ln(Rev)$			
	Dairy	Pig	Mixed	Crop
$\ln(CH_4)$	0.677^{***}	0.310^{***}	0.042^{**}	0.035^{***}
	(0.023)	(0.020)	(0.020)	(0.010)
$\ln(N_2O)$	0.077^{***}	0.074^{**}	0.220^{***}	0.133^{***}
	(0.021)	(0.031)	(0.035)	(0.041)
$\ln(CO_2)$	0.227^{***}	0.151^{***}	0.264^{***}	0.314^{***}
	(0.008)	(0.015)	(0.023)	(0.038)
$.5 \times \ln(CH_4)^2$	0.551^{***}	0.056^{***}	-0.035^{**}	0.005
	(0.141)	(0.017)	(0.015)	(0.004)
$.5 \times \ln(N_2O)^2$	0.720^{***}	-0.016	0.0004	0.105^{***}
	(0.152)	(0.093)	(0.082)	(0.024)
$.5 \times \ln(CO_2)^2$	0.185^{***}	0.079^{***}	0.093^{*}	0.334^{***}
	(0.021)	(0.026)	(0.055)	(0.040)
$\ln(CH_4) \times \ln(N_2O)$	-0.584^{***}	0.013	0.020	-0.010
	(0.134)	(0.042)	(0.024)	(0.007)
$\ln(CH_4) \times \ln(CO_2)$	0.032	-0.035	0.016	-0.004
	(0.043)	(0.029)	(0.018)	(0.008)
$\ln(N_2O) \times \log(CO_2)$	-0.132^{***}	0.043	-0.022	-0.189^{***}
	(0.049)	(0.049)	(0.062)	(0.028)
$Time$	0.005^{***}	0.022^{***}	0.0003	0.005
	(0.001)	(0.001)	(0.002)	(0.004)
$.5 \times Time^2$	0.001^{**}	0.006^{***}	0.004^{***}	0.009^{***}
	(0.0004)	(0.001)	(0.001)	(0.001)
$\ln(CH_4) \times Time$	-0.018^{***}	0.004	0.0001	0.0002
	(0.004)	(0.003)	(0.002)	(0.001)
$\ln(N_2O) \times Time$	0.017^{***}	-0.001	-0.011^{*}	0.005
	(0.004)	(0.005)	(0.006)	(0.004)
$\ln(CO_2) \times Time$	-0.002	0.003	0.015^{***}	0.002
	(0.002)	(0.003)	(0.005)	(0.005)
Observations	9,412	3,687	2,475	5,139
R^2	0.509	0.371	0.236	0.161
F Statistic	586.046^{***}	130.616^{***}	39.039^{***}	57.642^{***}

Note: $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

Table 3.5: KLM step 2 result table - estimation of time-varying emission inefficiency (bootstrapped standard errors, R = 1,000).

	Dairy		Pig		Mixed		Crop	
$Y = \hat{\epsilon}_{kt}$	Estimate	SE	Estimate	SE	Estimate	SE	Estimate	SE
Const.	0.073	(0.004)	0.104	(0.0059)	0.083	(0.0096)	0.211	(0.015)
σ^2	0.017	(< 0.001)	0.03	(0.0015)	0.025	(0.0014)	0.119	(0.0083)
γ	0.502	(0.0383)	0.575	(0.0406)	0.438	(0.0706)	0.611	(0.0574)

Table 3.6: KLM step 3 result table - estimation of time-invariant emission inefficiency (bootstrapped standard errors, R = 1,000).

	Dairy		Pig		Mixed		Crop	
$Y = \hat{\alpha}_i$	Estimate	SE	Estimate	SE	Estimate	SE	Estimate	SE
Const.	5.275	(2.533)	17.025	(1.5973)	11.564	(2.6558)	8.878	(4.222)
σ^2	0.056	(0.0022)	0.59	(0.04)	0.46	(0.062)	0.796	(0.0953)
γ	0.842	(0.0088)	0.907	(0.0098)	0.687	(0.0577)	0.715	(0.0458)

metrics, we need to check if revenue is monotonically increasing in GHG emissions. It can be seen from Table 3.4 that all GHGs are monotonically increasing in revenues at the sample mean. Table 3.7 summarizes the results of the monotonicity checks at all data points. Mixed farms show very few monotonicity violations: 0.85% out of all observations. Furthermore, the monotonicity condition is violated for 8.25% of observed crop farms and for 3.58% of pig farms. As for dairy farms, monotonicity violations are found for 27.25% of all observations and is observed in particular for N_2O, where revenue decreases in 26.54% of the observations.[13] In accordance with Sauer et al. (2006), for further analysis, we have dropped all observations that violate the monotonicity condition and keep only those observations that are theoretically consistent. [14]

3.6.2 Returns to emission scale

This section presents the results for the RTES. Kernel density plots of farm type specific RTES values are provided in Figure 3.1. This metric can be interpreted as

[13]As stated above, we have evaluated several alternative model specifications. Beside having been rejected by standard statistical tests, they performed worse than our chosen specification with respect to theoretical consistency.

[14]We avoid imposing monotonicity globally in order to retain functional flexibility in line with Sauer et al. (2006, p.161), who state *"[. . .]if theoretical consistency holds for a range of observations, this 'consistency area' of the estimated frontier should be determined and clearly stated to the reader. Estimated relative efficiency scores hence only hold for observations which are part of this range."*

Table 3.7: Percentage (%) of monotonicity violations by greenhouse gas and farm type.

	Dairy	Pig	Mixed	Crop
CH_4	0	0.16	0.32	4.11
N_2O	26.54	2.74	0	0.86
CO_2	0.84	2.55	0.53	3.37
Total	27.25	3.58	0.85	8.25

the percentage change in revenues that is associated with a one percent increase (decrease) in total greenhouse gases. We find that almost all farms in the sample reveal decreasing RTES. However, 30% of dairies show increasing RTES. Mean RTES are 0.97, 0.49, 0.50, 0.51 for dairy, pig, mixed and crop farms, respectively. Hence, if a crop farm decreases its emissions by 1%, revenues will decrease by 0.51%, which is, as before, an underproportional decline in revenues. Taking on

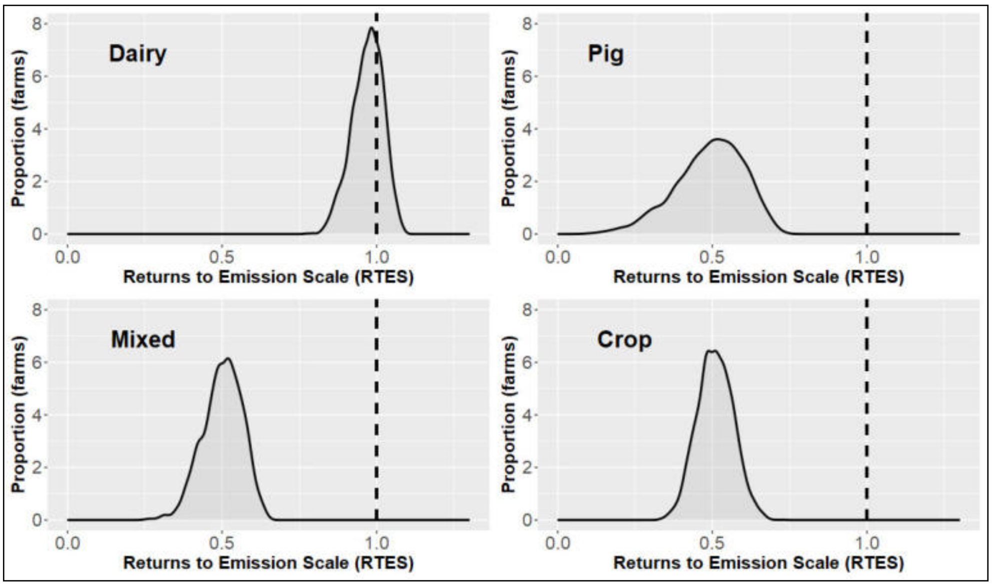


Figure 3.1: Returns to emission scale distributions of different farm types.

a social planner perspective and giving economic performance (in the form of revenues) and GHG emissions equal weights, constant RTES would be desirable. In that case emissions and revenues would be proportionate and a one percent increase (decrease) in emissions would be associated with a one percent increase (decrease) in revenues. From that perspective, dairy farms are on average close to the optimal emission scale, while the other farm types are far away from this point.

In allowing returns to emission/pressure scale to be variable, we can actually check whether or not the prevalent assumption of constant returns to scale as implied by most previous studies is plausible. Our results indicate that this is not the case for most Bavarian farms. Given similar production conditions, we can assume this holds also for other European regions. Whitney-Mann tests for all farming systems reject the null hypothesis that RTES are on average equal to one at standard confidence levels. However, evaluated at the sample mean, the 95% as well as the 99% confidence interval for the RTES of dairy farms contain the one (Annex Table 3.10). Also, 90.6% of the observations in the dairy subsample exhibit RTES of between 0.9 and 1.1. Hence, assuming constant RTES appears rather plausible for dairy farms. All other farm types exhibit strongly decreasing RTES – also when evaluated at the sample mean.

According to our models, nearly all farms reveal quite extreme RTES and are a long way away from the optimal emission scale size from an ecological-economic perspective. A reason for this finding could be the fact that farms do not seek to optimize their scale size with respect to emissions but rather with respect to input use in order to optimize their economic performance, rather than their ecological performance. Hence, farms may be at their most productive (economically motivated) scale size which is desirable from a manager's perspective. If their input usage is linked to high levels of emissions, they might reveal decreasing RTES and are far away from the societally desirable, most productive emission scale size, which would imply fewer GHG emissions.

3.6.3 Emission efficiency

Following Kumbhakar et al. (2014), time-varying emission efficiency scores were computed based on (3.19) while permanent efficiencies were calculated based on Eq. 3.20. The product for these two metrics equals total emission efficiency. Figure 3.2 presents the kernel density distributions of the individual efficiency scores for all farm types. Time-varying efficiency is very high across all farm types. Average scores range from 0.82 for crop farms to 0.92 for mixed farms and 0.93 for dairy farms (pig farms 0.9). Besides the higher mean scores, the spread for livestock-keeping farm typologies is smaller than the spread of crop farms.

Residual inefficiency stems from short-term rigidities on the farms (Kumbhakar et al., 2014). As livestock farmers are usually subject to a rather fixed environment in the form of stables, there is little room for improving emission efficiency in the short-run through improved management. The emission efficiency of crop farms, on the other hand, can be influenced by a multitude of short-term managerial decisions. For instance, the timing of certain activities plays an important role in terms of what yield can be obtained from a fixed set of inputs (International Fertilizer Industry Association (IFA), 1992). Thus, if inputs such as nitrogen fertilizers are applied in a timely fashion, inevitable N_2O emissions from this input use can be better translated into output, and from that, eventually into revenue. As for time-invariant efficiency, individual scores are on average lower than for

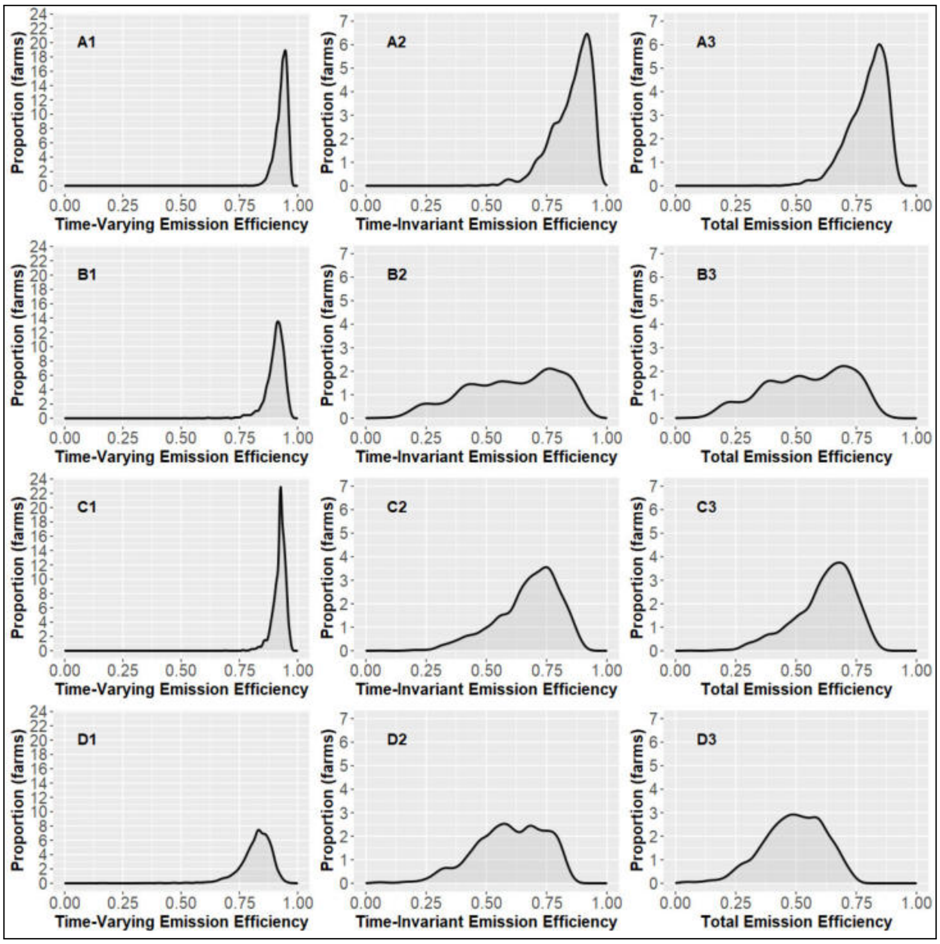


Figure 3.2: Density plots of farms' residual (1), permanent (2) and total emission efficiency (3) - Dairy (A), Pig (B), Mixed (C), Crop (D).

residual efficiency. This is particularly true for pig, crop and mixed farms. The distributions are rather wide for pig, mixed and crop farms, i.e. there is greater heterogeneity across farms than for dairy farms. The fact that persistent efficiency is rather low could be indicative that farmers need to make structural changes to improve overall efficiency through e.g. farm size adjustments or investments in input-efficient, climate-friendly technologies. The high structural inefficiency may as well explain why on average the farmers operate under decreasing economies of scale.

Finally, overall efficiency is on average highest for dairy farms (79.5%), followed by mixed farms (62.0%), pig farms (54.9%) and crop farms (49.1%).[15] This means that, say, pig farms only generate 54.9% of their maximum revenue given their level of damage on the climate. This also means by implication that farmers could considerably reduce climatic stress while maintaining the same level of revenue. This finding is in line with most previous studies on eco-efficiency (e.g. Godoy-Durán et al., 2017; Picazo-Tadeo et al., 2012, 2011). As for crop farms, Bonfiglio et al. (2017) and Gadanakis et al. (2015) find mean eco-efficiency scores of 54.8% and 56.2% in Italy and the UK, respectively. Pérez Urdiales et al. (2016) and Orea & Wall (2017) report average eco-efficiency levels for dairy farms in Asturia (Spain), which are markedly lower than the average in our sample. With regard to pig and mixed farms, we could not find any comparable study.

The huge potential for reducing climatic damage without affecting economic performance is particularly striking. From a societal point of view it is important to ponder potential reasons for the high level of eco-inefficiency found in our analysis. One reason could lie in the technical inefficiency of farms (Picazo-Tadeo et al., 2011). Previous studies found a strong relationship between TE and eco-efficiency (Beltrán-Esteve et al., 2014; Gadanakis et al., 2015; Picazo-Tadeo et al., 2011). If farmers manage their inputs efficiently such that they can reduce their level of input use while maintaining their output level, then they are also likely to be eco-efficient. Conversely, the overuse of inputs such as nitrogen leads to technical and ecological inefficiency. Picazo-Tadeo & Reig-Martínez (2006) show in their study how pressures on the environment could be reduced by simply promoting best farming practices. This principle applies to both time-varying and persistent inefficiency. Differences in TEs may also serve as an explanation for emission efficiency differences in our case study. Mennig & Sauer (2020) find higher average TE scores for dairy farms than for crop farms in Bavaria between 2007 and 2011, which corresponds to our findings on eco-efficiency.

Furthermore, various other aspects have been found in the literature to have an effect on eco-efficiency. Pérez Urdiales et al. (2016) find that age has a negative effect on eco-efficiency, i.e. older farmers are less eco-efficient, which is also shown by Reinhard et al. (2002). Another reason for varying eco-efficiency scores

[15] Efficiency levels are computed for the respective farm typology-specific technologies. Hence, we investigate efficiency levels relative to the farm type specific benchmarks and not for an overall meta-technology.

could relate to farmers' education level, which is assumed to be closely linked to managerial skills. According to Picazo-Tadeo et al. (2011) and Gadanakis et al. (2015) a higher education level is positively associated with eco-efficiency. Also, the prospect of farm succession seems to play an important role with respect to the level of eco-efficiency. Pérez Urdiales et al. (2016) and Bonfiglio et al. (2017) show that if there is a positive expectation of the farm continuing, farms are more eco-efficient. Furthermore, policy interventions such as agri-environment schemes (AES) or stricter environmental regulations have been found to be positively associated with the eco-efficiency of arable farms (Bonfiglio et al., 2017; Gadanakis et al., 2015; Pérez Urdiales et al., 2016).

3.6.4 Eco-performance dynamics

So far, we have only considered farms' eco-performance from a static point of view. In this section, we seek to investigate the dynamic structure of eco-performance and its components. As outlined in Section 3.5.2, eco-performance dynamics are determined by emission scale change, technical change and emission efficiency change.[16] Ultimately, eco-performance can be viewed as the synthesis of the concepts presented in the previous sections.

Mean change rates of eco-performance and its components for all four subsamples from 2005-2014 are presented in Table 3.8. Annual change rates are 0.49% for dairy farms, 1.97% for pig farms, 0.08% for mixed farms and -0.04% for crop farms. Eco-performance growth was mainly due to technical progress with respect to emissions for dairy and pig farms. The fact that technical change is the main driver of eco-productivity has previously been shown in the literature (Beltrán-Esteve & Picazo-Tadeo, 2017; Kortelainen, 2008; Picazo-Tadeo et al., 2014). Emission-efficiency change is shown to be the highest (1.09%) for crop farms. The other farm types reveal efficiency change rates rather close to zero. Average emission scale change rates are close to zero for all industries other than crop farms, where a negative annual development of, on average, -1.1% has been identified.

Regarding technical change with respect to emissions, all industries reveal a positive trend in change rates starting at different points in time. Pig farms, mixed farms, and crop farms experienced technical progress in the period under review as of 06/07, 09/10 and 10/11, respectively. One reason for the technical regres-

[16]Note, that the EEC only represents the transient component of inefficiency as persistent inefficiency is stable over time.

Table 3.8: Emission performance change (EPC) decomposed into scale change (SEC), ecological-technical change (ETC), and efficiency change (EEC) expressed as percentage changes (%).

	Dairy				Pig				Mixed				Crop			
Period	SEC	ETC	EEC	EPC	SEC	ETC	EEC	EPC	SEC	ETC	EEC	EPC	SEC	ETC	EEC	EPC
05/06	0.02	0.41	-0.35	0.07	-0.26	-0.24	-0.96	-1.46	-0.14	-1.82	-1.43	-3.39	-1.14	-3.71	-5.19	-10.04
06/07	0.02	0.42	-4.64	-4.19	-0.59	0.33	1.34	1.09	0.01	-1.37	-1.76	-3.12	-1.12	-2.85	-4.27	-8.25
07/08	0.02	0.48	4.92	5.43	-0.33	0.94	-2.25	-1.64	0.14	-1.06	3.18	2.25	-1.07	-1.93	14.75	11.74
08/09	0.09	0.55	1.37	2.02	-0.57	1.57	4.32	5.32	-0.03	-0.37	0.92	0.51	-1.76	-0.96	2.85	0.13
09/10	0.05	0.54	-3.99	-3.40	-0.82	2.13	-0.11	1.20	0.36	-0.00	-1.93	-1.58	-1.28	-0.06	-4.33	-5.67
10/11	0.01	0.61	0.43	1.06	-0.46	2.71	-0.79	1.45	0.39	0.29	-0.17	0.51	-1.12	0.81	0.18	-0.13
11/12	0.04	0.71	2.15	2.90	-0.25	3.31	-1.93	1.13	0.36	0.68	-0.93	0.12	-1.03	1.76	-0.59	0.14
12/13	0.02	0.73	-3.19	-2.44	-0.31	3.97	1.87	5.53	0.16	1.23	1.29	2.69	-0.78	2.84	6.77	8.82
13/14	0.04	0.81	2.16	3.01	-0.09	4.59	0.62	5.13	0.32	1.84	0.57	2.73	-0.60	3.85	-0.38	2.87
per annum	0.04	0.58	-0.12	0.49	-0.41	2.15	0.23	1.97	0.18	-0.06	-0.03	0.08	-1.10	-0.03	1.09	-0.04

sion in the pressure-generating technology could be that the underlying *production technology* may have altered such that the input combination of farmers leaned toward more emission-intensive inputs. As concerns the emission efficiency change, sharp increases in efficiency are found for the period 2007/2008 for dairy, mixed and crop farms and for 2008/2009 with respect to pig farms. This is followed by periods of decreasing growth, efficiency decay and recovery at different rates and following different patterns later on.

Figure 3.3 depicts the mean, first and third quantile values of the composite eco-performance patterns between 2005 and 2014. As before, we can see different patterns for different industries. The smallest average degree of volatility was found in mixed farms, while crop farms were characterized by high fluctuations. Dairy and pig farms were found to be somewhat between the two extremes. Hence, the above-mentioned overall eco-performance improvement did not develop in a monotonically increasing fashion for any of the analyzed farm types. Additionally, the distribution of change rates expressed as the interquartile range is highest for crop farms, indicating considerable intra-industry eco-performance differences. That indicator is at its lowest for dairy farms, i.e. less intra-industry differences can be observed with respect to emission performance.

Eco-efficiency change is the key factor behind the eco-performance fluctuation patterns. As mentioned previously, one major factor that influences time-varying efficiency is the managerial ability of farmers regarding their input use. This gives rise to the essential question as to which contexts are farmers producing more efficiently in, compared to others. Figure 3.4 possibly delivers an explanation for that phenomenon. Taking the case of crop farms, we observe the real price development of cereal prices in the relevant time period (solid line in Figure 3.4). Low-price periods are followed by a decline in the eco-performance of crop farms (compare Figure 3.3), while high-price periods are usually followed by increasing rates in eco-productivity. Hence, if high output prices are expected, farmers could seek to manage their constrained inputs more efficiently than in times of low-price expectations. Similar but less pronounced movements can also be found for milk prices and eco-productivity of dairy farms as well as for pig prices, and for the eco-performance movements of pig fattening farms. Finally, assuming a close relationship between eco-performance and TFP, Frick & Sauer (2018) and Mennig & Sauer (2020) find similar TFP patterns for dairy and crop farms in Bavaria.

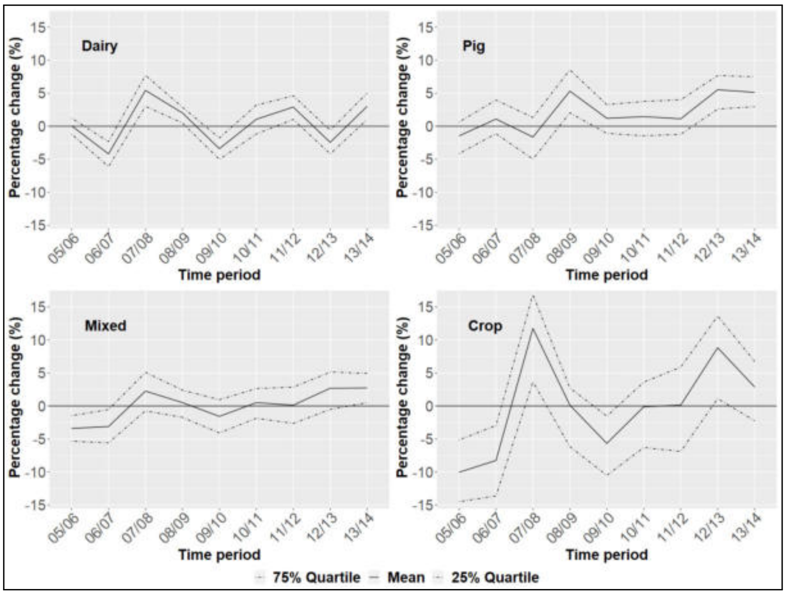


Figure 3.3: The mean, first and third quartile values (middle, bottom and top lines) of emission productivity of Bavarian sample farms.

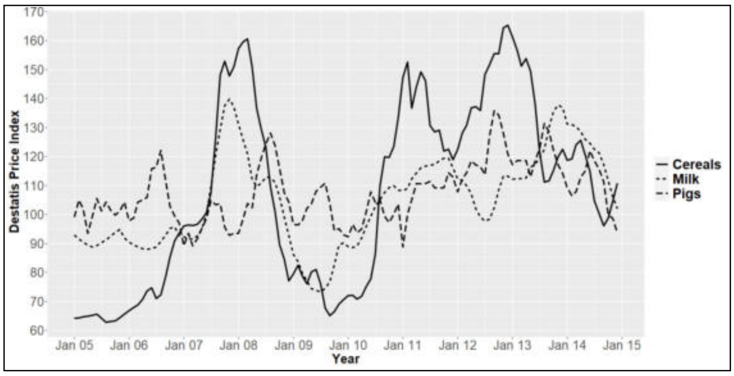


Figure 3.4: Destatis price index from January 2005 to December 2014 (monthly): cereals, milk, pigs.

3.6.5 Policy implications

The fact that most farms in the sample show strongly decreasing RTES raises the question as to what legislators can do if their objective is to observe a proportional relationship between emissions and revenue, i.e. if emissions are changed by one per cent, revenues change accordingly by one per cent. One obvious answer would be to regulate farm size to balance GHG emissions and economic outcome (input scale effect). However, this approach could have several consequences, e.g. TFP could decrease as farms' TE might decrease due to impaired input use management or underutilized resources. Hence, (regional) added value in agriculture could decline, which might have unintended consequences for parts of the rural population. Another approach is to foster policies that aim at decreasing the amount of emissions per unit of input (input management effect). This would allow farmers to remain at a productive scale in terms of the input-output relationship and at the same time move towards the most eco-productive scale. For instance, legislators could promote the use of precision agriculture practices for crop producers such as global positioning systems, where input application and agronomic practices are matched with soil attributes (Gadanakis et al., 2015). As for livestock farms, manure management systems could be improved to reduce GHG emissions per livestock unit. For instance, Petersen et al. (2013) find that covering up manure storage facilities or treating manure with additives can substantially decrease CH_4 releases. These measures are expected to also positively affect persistent emission efficency of farms through investments leading to (eco-)structural changes.

As for farmers' general eco-performance with respect to greenhouse gas emissions, there are ample options available to policy-makers. First, legislators could promote agricultural training programs aiming at improving farmers' managerial skills, which eventually translates into improved input-management and better emission efficiency. Picazo-Tadeo et al. (2012, p.806) note that policies aiming at increasing productive efficiency "*[...] can be considered the most cost-efficient way of reducing environmental pressures without reducing farmers' income*". Legislators should thereby take farm type specificities into account as performance varies strongly across farm types.

Second, various policy options exist that aim to internalize environmental externalities induced by farming activities. This also applies to the emission of greenhouse gases. By more effectively conditioning farmers' income to their climate-protection performance, a behavior which is more oriented towards the public good can be expected (compare Beltrán-Esteve et al., 2014; Picazo-Tadeo et al., 2012). E.g. Picazo-Tadeo et al. (2011) demand a stronger commitment of EU policy-makers to the principle of *conditionality*, i.e. only farmers that comply with ambitious ecological standards should benefit from public resources. Moreover, farmers could be sanctioned for adverse climatic performance (*polluter-pays* principle) or could be financially rewarded for climate-friendly farming practices (*provider-gets* principle). The most pressing need for action applies to those farm types that were found to be on average very emission-inefficient, such as crop and pig farms. Furthermore, EU second-pillar AES are considered to promote eco-efficiency. However, several authors note the cost-inefficiency of such AES (Beltrán-Esteve et al., 2014; Bonfiglio et al., 2017; Picazo-Tadeo et al., 2011).

Finally, policy-makers are not overly concerned about short-run fluctuations in emission-performance. Since the ultimate objective is to mitigate climate change and its adverse effects on the environment, a positive long-run development of eco-performance and its determinants is pivotal to that end. Beside the aforementioned instruments to improve RTES and eco-efficiency, ecological-technical progress could be promoted. Beside the aforementioned adoption of existing climate-friendly technologies, legislators could stimulate eco-innovations in the context of climate-smart agriculture. Long et al. (2016) recommend, among other things, financial support for start-up companies and tax-cuts for research and development activities. This could boost technological improvements and have a positive impact on farms' emission-performance and finally on their relative climate change mitigation potential.

3.7 Summary and concluding remarks

Increasing concerns over the environmental implications of farming activities as well as the need for increasing productivity require the development of monitoring and evaluating instruments with respect to the ecological-economic performance of farm businesses, in short "eco-performance". This is particularly true with respect to the trade-off between climatic impact and the economic performance of farms as the adverse effects of global warming become ever more apparent. Measuring eco-performance in terms of GHG emissions is important, as it might provide policy-makers and farm managers with sound information for designing measures to reduce GHG emissions while at the same time improving economic performance.

This paper has presented an approach to assess firms' relative climate change mitigation potential by building upon and further developing the concept of eco-efficiency (Kuosmanen & Kortelainen, 2005; Orea & Wall, 2017). We presented a parametric stochastic frontier approach capable of capturing eco-performance dynamics over time. Unlike previous studies on eco-efficiency, we allowed for a complex functional form to aggregate ecological (climatic) pressures into environmental damage. The resulting *'pressure conversion function'* describes how well ecological pressures translate into economic output. Moreover, we considered the fact that the underlying pressure-generating technology might exhibit variable returns to pressure scale. Finally, our theoretical framework let us analyze eco-performance dynamics and its components – technical change, scale change and eco-efficiency change by means of a generalized Malmquist productivity index.

Our empirical application focused on four different farm types in the German federal state of Bavaria, which is a significant and somewhat representative region in the EU for the years 2005 to 2014. In order to approximate GHG emissions at farm-level, we make use of recent methodological developments in the literature by applying and downscaling the IPCC guidelines for national greenhouse gas inventories to the firm-level. Besides farm accountancy data, a unique combination of various data sources is used to estimate the pressure-generating technology separately for dairy, pig, mixed and crop farms, based on a stochastic frontier model for panel data based distinguishing between time-varying and persistent eco-inefficiency.

The main findings of the study were the following. Our study revealed that, evaluated at the sample mean, all greenhouse gases (namely CH_4, N_2O and CO_2) were positively associated with farms' revenues. Mann-Whitney tests reject the hypothesis of constant returns to scale for all farm types. Nearly all farms in our sample experience strongly decreasing returns to their scale of GHG emissions. However, 90% of dairy farms exhibit emission scale elasticities of between 0.9 and 1.1, i.e. they are close to the optimal emission scale of one. As for emission efficiencies, farms revealed little time-varying eco-inefficiency and rather high levels of persistent inefficiency. Overall, the farms in our sample were very eco-inefficient. Dairy farms were on average least eco-inefficient (∼80%) followed by mixed (∼60%) and pig farms (∼55%) and crop farms (∼50%). Based on the two above-mentioned criteria, RTES and emission-efficiency, dairy farms performed best, followed by pig farms, mixed farms and finally crop farms. In terms of eco-performance dynamics, our results showed that pig farms revealed the highest annual growth rates followed by dairy and mixed farms. Crop farms exhibited very little (but slightly negative) eco-performance growth between 2005 and 2014.

As this study presents a novel approach for measuring eco-performance at the micro-level, it is subject to several limitations. First, our method for retrieving GHG emissions at the micro-level represents an approximation of the true GHG releases. By combining data from different aggregation levels and sources, some imprecisions are inevitable. Also, our method only accounts for GHG releases. We do not consider possible farm management options that serve as carbon sinks. Second, our eco-performance measure does not imply total ecological pressures. Instead, we focus on climatic stress. Therefore, we are not able to capture potentially important interactions between climatic stress and other ecological stresses such as nitrogen surpluses or pesticide risk. Third, some caution is due when interpreting our estimation results, especially with respect to dairy farms. As there is a non-negligible proportion of farms that do not comply with the monotonicity condition, we have only made statements on theoretically consistent observations. Finally, it should be noted that the presented concept does not refer to farms' absolute ecological performance, but relative eco-performance with respect to GHG emissions. Even if farms were completely eco-efficient, absolute emission levels could still be too high, in that the carrying capacity of the atmosphere is exceeded.

Several policy options for improving farms' emission-performance can be suggested: Promoting the development of eco-innovations and climate-smart technologies as well as fostering the adoption of already existing climate-smart technologies might be one way forward. Also, the EU's common agricultural policy could be more strongly oriented toward the principle of conditionality, i.e. only farmers that comply with high emission standards would profit from public support. Other options include training programs to improve farmers' managerial skills and raise environmental awareness. Nevertheless, even if farmers exploit their full potential regarding emission performance, more drastic policy measures to reduce absolute levels of GHG released might be necessary to further limit climate change.

Finally, we would like to highlight potential avenues for further research. From a methodological point of view, it would be interesting to directly incorporate eco-efficiency determinants into the error term of the pressure conversion frontier. This would allow for the evaluation of their impact on eco-efficiency. Moreover, it would be interesting make comparisons across pressure-generating technologies by using a metafrontier framework. Furthermore, the consideration of more environmental pressures than GHG emissions would be insightful from a more general sustainability perspective, as it would provide a more holistic picture of the ecology-economy trade-off. From an empirical standpoint, data on environmental pressures can be rare and difficult to obtain. It would be advisable to enrich, e.g., farm accountancy data with additional information on environmental pressures. This would also be helpful in further improving the accuracy of the farm-level greenhouse-gas-inventory methodology presented in this paper. Another interesting avenue for future research could be the evaluation of the econnection between the environmental Kuznet curve and eco-efficiency at the micro-level. Lastly, it would also be insightful to extend the eco-performance concept by incorporating factors representing the social dimension of a more holistic sustainability concept.

3.8 Appendices

3.8.1 Data processing flow to obtain farm-level GHG emissions

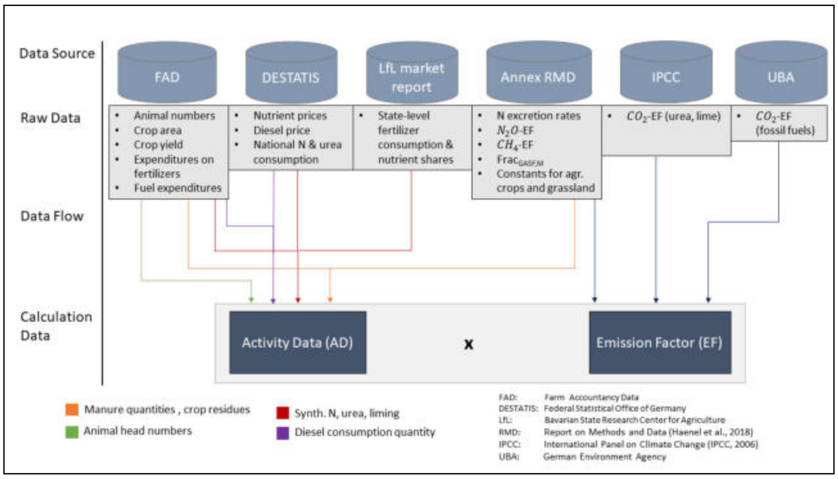


Figure 3.5: Data processing flow. Several data sources are considered to obtain farm-level GHG emissions.

3.8.2 Description of GHG emission sources within the agricultural context

Animal husbandry and enteric fermentation. Herbivores, ruminants in particular, produce CH_4 as a by-product of a digestive process whereby carbohydrates are broken down by microorganisms. Haenel et al. (2018) provides regional (state-level) EFs for the above-mentioned livestock categories which are multiplied by the number of animals stemming from the accountancy data.

Manure Management. Direct GHG emissions come from the storage and handling of livestock excrement. Anaerobic processes cause the formation of methane; nitrification and denitrification cause N_2O emissions. Excreted nitrogen quantities are calculated by using regional excretion factors per livestock unit provided by Haenel et al. (2018). Different manure storage systems yield different GHG emission quantities. Here, due to data restrictions from the accountancy data, average EFs (still Bavaria-specific) are applied. What is more, there are also indirect N_2O emissions coming from deposition of reactive nitrogen (i.e. ammonia and nitric oxide). Again, these are accounted for by way of respective EFs.

Fertilizer application. In agricultural soils, N_2O is produced as a by-product of nitrification and denitrification processes and leaks into the atmosphere. Adding nitrogen through the application of fertilizers (both synthetic and organic) thus generates N_2O emissions. There are year and region specific EFs for organic and synthetic N quantities available. However, the existing accountancy data solely delivers overall fertilizer expenditures per year. [17] Another source of N_2O emissions connected to fertilizers is leaching and surface run-off of N. A fixed fraction of applied N (0.3,IPCC (compare 2006)) is assumed to be lost through this pathway which brings about additional nitrous oxide emissions (Haenel et al., 2018; IPCC, 2006). The IPCC default EFs is used. CO_2 emissions from the application of urea is determined analogous to the direct N_2O emissions from N-fertilizer use. The use of calcium fertilizers to reduce the acidity of soils (liming) causes the emission of CO_2.

Crop residues. Crop residues left on agricultural fields contain nitrogen. Due to microbial processes of theses residues in the ground, N_2O emissions are released into the atmosphere. A fine distinction is made with respect to crop/plant type, and farm-level yield data is used to recover the precise amount of emissions originating from crop residues.

Energy use. To approximate emissions from the on-farm use of energy, we calculate the yearly diesel quantities by using official diesel price data as well as diesel expenditure from the accountancy data. In accordance with the IPCC (2006) instructions, and using the regional EFs provided by the Federal Environment Agency, CO_2 emissions are calculated.

Other sources. Other sources of GHG emissions from agriculture include rice cultivation (methane) and the burning of crop residues and stubble fields (CO_2). While rice is not cultivated in Germany, field burning is negligible as it is prohibited under law.

3.8.3 Four-step procedure for approximating fertilizer quantities and plausibility checks

Four-step procedure

- First, based on the shares of pure nutrients in total fertilizers, the yearly N share is calculated for Bavaria.
- Second, yearly pure nutrient prices are multiplied by these shares to obtain a price index for pure nutrients.

[17] More recently, farmers within the FADN framework have been obliged to report more detailed information as to what fertilizers are utilized.

- Third, farm-level fertilizer expenditures are divided by the price index to obtain fertilizer quantities expressed in terms of pure nutrients.
- Fourth, the average share of N in total fertilizers for all farms is assumed, so as to finally obtain the nitrogen quantity from synthetic fertilizers.[18]

Plausibility checks of the approximated nitrogen quantities

First, in order to assess if our 4-step procedure for approximating farm-level nitrogen (N) application is plausible, we detected observations with unreasonably high amounts of N fertilizer. Only 0.35% of all observations apply more than 300 kg N per ha. Also, according to our approach 3.92% of all observation use is between 200 and 300 kg N per ha. Hence, more than 95% of the data have a synthetic N usage of up to 200 kg/ha, which appears to be quite plausible.

Furthermore, we compared two cases of total farm-level GHG emissions expressed as CO_2-equivalents. In one approach, we utilized N quantities calculated using our 4-step procedure. In the other case, we computed fertilizer-related emissions based on the mean nitrogen per hectare of cropland or grassland, as provided by the Bavarian State Research Center for Agriculture (LfL). A summary of the results is provided in Table 3.9. There is a negligible difference between the two methods for dairy farms. For the other farm types, the 4-step procedure yields markedly higher levels of GHG emissions. Although we do not know the true amount of emissions, our approach might lead to an overestimation of actual farm-level GHG emissions. Nevertheless, we adhere to the suggested 4-step-procedure, as it returns sensible differences in the application of synthetic N fertilizers across farm types, which would be neglected if we used average N/ha amounts for all farms. The left panel of Figure 3.6 shows different synthetic N application rates for crop and livestock farms. As expected, livestock farms lie well below crop farms as they presumably cover their nitrogen fertilizer need partly through farm manure. A similar picture is shown in the right panel of Figure 3.6. Crop farms exhibit the highest need for synthetic nitrogen, followed by pig and mixed farms. Surprisingly, dairy farms lie well below the other farm types. This might be due to the fact that dairies in Bavaria are traditionally grassland-based and might consequently need less synthetic nitrogen.

[18]***Note:*** Within the FADN, most countries have recently started to collect data on fertilizer quantities, which will allow for a more precise reconstruction of fertilizer-related GHG emissions.

Table 3.9: Summary of farm-level GHG emissions when applying different methods for approximating synthetic nitrogen usage. Results based on the 4-step procedure described in Section 3.4.2 are compared to the results based on the average per ha synth. N quantity in Bavaria as provided by the Bavarian Research Center for Agriculture (LfL).

	Synth. N calculation method	Min. (t CO_{2e})	1st Qu. (t CO_{2e})	Median (t CO_{2e})	Mean (t CO_{2e})	3rd Qu. (t CO_{2e})	Max. (t CO_{2e})
Dairy	4-step N computation	79.89	363.4	483.78	546.52	655.22	2173.36
	LfL mean kg N per ha	82.7	365.39	485.15	546.82	653.73	2180.04
Pig	4-step N computation	18.81	137.67	243.15	272.58	368.41	1058.02
	LfL mean kg N per ha	15.36	131.53	232.91	259.51	358.58	996.74
Mixed	4-step N computation	24.17	135.42	247.23	302.23	409.55	1282.41
	LfL mean kg N per ha	22.86	121.23	230.35	282.29	383.25	1172.55
Crop	4-step N computation	6.56	47.63	80.39	116.09	136.99	726.15
	LfL mean kg N per ha	9.41	36.9	60.69	93.07	101.89	643.4

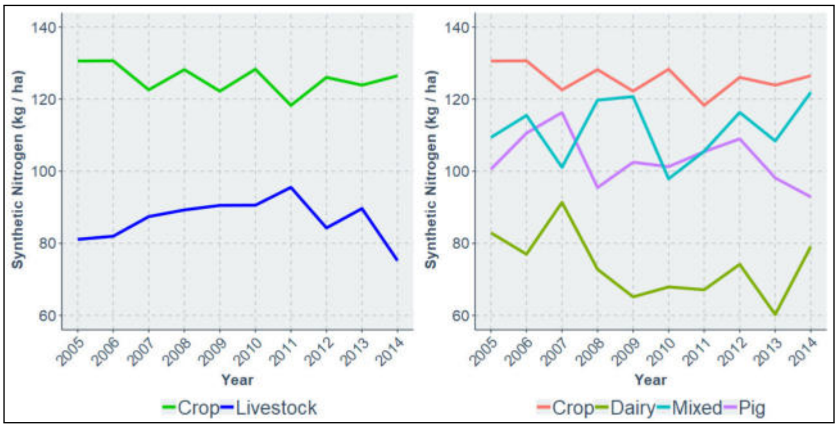


Figure 3.6: Synthetic nitrogen use per hectare in accordance with suggested 4-step-procedure.

3.8.4 Three-step estimation approach of the KLM stochastic frontier model

Starting from this formulation Kumbhakar et al. (2014) propose the following 3-step approach:

***Step 1*:** A standard fixed-effects regression is employed to estimate the parameters ($\hat{\mathbf{fi}}$) in (3.11). This allows us to compute parameter-based measures such as partial elasticities and returns to scale in the conventional fashion. Step 1 also delivers predicted values of α_k and ϵ_{kt}, denoted $\hat{\alpha_k}$ and $\hat{\epsilon}_{kt}$. To assess the robustness of our estimator, we test it against the pooled regression model.

Furthermore, our model is tested against the more consistent random-effects estimator.

Step 2: Time-varying emission inefficiency, u_{kt}, is estimated by making use of $\hat{\epsilon}_{kt}$ and further distributional assumptions. We know that

$$\epsilon_{kt} = \nu_{kt} - u_{kt} + E(u_{kt}). \tag{3.19}$$

Inserting the predicted value $\hat{\epsilon}_{it}$ into the LHS of the equation and assuming for $\nu_{kt} \sim N(0, \sigma_\nu^2)$ and $u_{kt} \sim N^+(0, \sigma_u^2)$, such that $E(u_{kt}) = \sqrt{2/\pi\sigma_u}$, we can estimate (3.19) by the standard stochastic frontier technique. Time-varying emission efficiency (REE) is obtained through $e^{(-u_{kt}|\epsilon_{kt})}$ (Battese & Coelli, 1988).

Step 3: Time-invariant emission efficiency is estimated similar to the approach in step 2. We know that

$$\alpha_k = \mu_k - \eta_k + E(\eta_k). \tag{3.20}$$

By assuming $\mu_{kt} \sim N(0, \sigma_\mu^2)$ and $\eta_{kt} \sim N^+(0, \sigma_{eta}^2)$, so that $E(\eta_{kt}) = \sqrt{2/\pi\sigma_\eta}$, and replacing α_k with $\hat{\alpha_k}$, we obtain estimates for persistent emission efficiency (PEE), $e^{-\eta_k}$.

Overall emission efficiency is the product of REE and PEE. As noted by Kumbhakar et al. (2014), step 2 and step 3 yield biased standard errors, as the independent variables are predicted values from step 1 and not the observed values. We use clustered bootstrapping to obtain correct standard errors for the respective coefficient estimates.

3.8.5 Bootstrapped confidence interval for returns to emission scale at the sample mean

Table 3.10: Bootstrapped confidence intervals (R=1000) for returns to emission scale (RTES) evaluated at the sample mean.

	Dairy		**Swine**		**Mixed**		**Crop**	
	L.B.	U.B.	L.B.	U.B.	L.B.	U.B.	L.B.	U.B.
95% Confidence Interval	0.964	1.031	0.485	0.572	0.473	0.581	0.437	0.586
99% Confidence Interval	0.953	1.042	0.471	0.586	0.456	0.598	0.413	0.609

L.B. = lower bound; U.B. = upper bound

3.8.6 Robustness checks w.r.t capital and labor

Table 3.11: Comparison of the original model and an alternative model including capital and labor as independent variables for dairy and pig fattening farms. The coefficient estimates for the alternative models are very similar to the original ones and there is no qualitative difference in these estimates. To further check for the occurence of omitted variable bias, we recalculated all results presented in Section 5 (i.e. elaticities, rtes, efficiencies etc). The results for the alternative specifications were virtually the same as for our original model. Hence, we conclude that our model does not suffer from omitted variable bias due to the omission of capital and labor.

	Dependent variable:			
	$\ln(Rev)$			
	Dairy orig.	Dairy altern.	Pig orig.	Pig altern.
$\ln(CH_4)$	0.678^{***}	0.669^{***}	0.310^{***}	0.305^{***}
	(0.023)	(0.023)	(0.020)	(0.020)
$\ln(N_2O)$	0.076^{***}	0.077^{***}	0.074^{**}	0.067^{**}
	(0.021)	(0.021)	(0.031)	(0.031)
$\ln(CO_2)$	0.226^{***}	0.225^{***}	0.151^{***}	0.147^{***}
	(0.008)	(0.008)	(0.015)	(0.015)
$.5 \times \ln(CH_4)^2$	0.555^{***}	0.550^{***}	0.056^{***}	0.057^{***}
	(0.141)	(0.141)	(0.017)	(0.017)
$.5 \times \ln(N_2O)^2$	0.732^{***}	0.733^{***}	-0.016	-0.005
	(0.154)	(0.154)	(0.093)	(0.093)
$.5 \times \ln(CO_2)^2$	0.186^{***}	0.185^{***}	0.079^{***}	0.079^{***}
	(0.021)	(0.021)	(0.026)	(0.026)
$\ln(CH_4) \times \ln(N_2O)$	-0.591^{***}	-0.590^{***}	0.013	0.007
	(0.135)	(0.135)	(0.042)	(0.042)
$\ln(CH_4) \times \ln(CO_2)$	0.036	0.037	-0.035	-0.034
	(0.044)	(0.044)	(0.029)	(0.029)
$\ln(N_2O) \times \log(CO_2)$	-0.137^{***}	-0.137^{***}	0.043	0.040
	(0.049)	(0.049)	(0.049)	(0.048)
$Time$	0.005^{***}	0.005^{***}	0.022^{***}	0.023^{***}
	(0.001)	(0.001)	(0.001)	(0.001)
$.5 \times Time^2$	0.001^{**}	0.001^{**}	0.006^{***}	0.006^{***}
	(0.0004)	(0.0004)	(0.001)	(0.001)
$\ln(CH_4) \times Time$	-0.018^{***}	-0.018^{***}	0.004	0.004
	(0.004)	(0.004)	(0.003)	(0.003)
$\ln(N_2O) \times Time$	0.017^{***}	0.017^{***}	-0.001	-0.002
	(0.004)	(0.004)	(0.005)	(0.005)
$\ln(CO_2) \times Time$	-0.002	-0.002	0.003	0.003
	(0.002)	(0.002)	(0.003)	(0.003)
$\ln(Labor)$		0.009		0.012
		(0.012)		(0.021)
$\ln(Capital)$		0.009^{*}		0.034^{***}
		(0.005)		(0.008)
Observations	9,412	9,412	3,687	3,687
R^2	0.509	0.510	0.371	0.375
F Statistic	586.015^{***}	513.057^{***}	130.616^{***}	116.117^{***}
Note:				*p<0.1; **p<0.05; ***p<0.01

Table 3.12: Comparison of the original model and an alternative model including capital and labor as independent variables for mixed and crop farms. The coefficient estimates for the alternative models are very similar to the original ones and there is no qualitative difference in these estimates. To further check for the occurence of omitted variable bias, we recalculated all results presented in Section 5 (i.e. elaticities, rtes, efficiencies etc). The results for the alternative specifications were virtually the same as for our original model. Hence, we conclude that our model does not suffer from omitted variable bias due to the omission of capital and labor.

	Dependent variable:			
	$\ln(Rev)$			
	Mixed orig.	Mixed altern.	Crop orig.	Crop altern.
$\ln(CH_4)$ [a]	0.042^{**}	0.040^{*}	0.035^{***}	0.037^{***}
	(0.020)	(0.020)	(0.010)	(0.010)
$\ln(N_2O)$	0.220^{***}	0.217^{***}	0.133^{***}	0.126^{***}
	(0.035)	(0.035)	(0.041)	(0.041)
$\ln(CO_2)$	0.264^{***}	0.262^{***}	0.314^{***}	0.310^{***}
	(0.023)	(0.023)	(0.038)	(0.038)
$.5 \times \ln(CH_4)^2$	-0.035^{**}	-0.035^{**}	0.005	0.006
	(0.015)	(0.015)	(0.004)	(0.004)
$.5 \times \ln(N_2O)^2$	0.0004	-0.009	0.105^{***}	0.103^{***}
	(0.082)	(0.082)	(0.024)	(0.024)
$.5 \times \ln(CO_2)^2$	0.093^{*}	0.091^{*}	0.334^{***}	0.329^{***}
	(0.055)	(0.055)	(0.040)	(0.040)
$\ln(CH_4) \times \ln(N_2O)$ [a]	0.020	0.020	-0.010	-0.010
	(0.024)	(0.024)	(0.007)	(0.007)
$\ln(CH_4) \times \ln(CO_2)$[a]	0.016	0.016	-0.004	-0.004
	(0.018)	(0.018)	(0.008)	(0.008)
$\ln(N_2O) \times \log(CO_2)$	-0.022	-0.016	-0.189^{***}	-0.186^{***}
	(0.062)	(0.062)	(0.028)	(0.028)
$Time$	0.0003	0.0003	0.005	0.006
	(0.002)	(0.002)	(0.004)	(0.004)
$.5 \times Time^2$	0.004^{***}	0.004^{***}	0.009^{***}	0.009^{***}
	(0.001)	(0.001)	(0.001)	(0.001)
$\ln(CH_4) \times Time$ [a]	0.0001	-0.0001	0.0002	0.0002
	(0.002)	(0.002)	(0.001)	(0.001)
$\ln(N_2O) \times Time$	-0.011^{*}	-0.011^{**}	0.005	0.006
	(0.006)	(0.006)	(0.004)	(0.004)
$\ln(CO_2) \times Time$	0.015^{***}	0.014^{***}	0.002	-0.001
	(0.005)	(0.005)	(0.005)	(0.005)
$\ln(Labor)$		0.030		0.034
		(0.029)		(0.026)
$\ln(Capital)$		0.028^{**}		0.031^{***}
		(0.012)		(0.011)
Observations	2,475	2,475	5,139	5,139
R^2	0.236	0.239	0.161	0.163
F Statistic	39.039^{***}	34.630^{***}	57.642^{***}	51.273^{***}

Note: *p<0.1; **p<0.05; ***p<0.01

4 Study II – Are intensive farms more emission-efficient? Evidence from German dairy farms

Disclaimer:

This version of the article has been accepted for publication, after peer review but is not the Version of Record and does not reflect post-acceptance improvements, or any corrections. This pre-print version of the article (Stetter, C., Wimmer, S. & Sauer, J.(2022). Are Intensive Farms More Emission-Efficient? Evidence From German Dairy Farms. Journal of Agricultural and Resource Economics) is available online at: `http://dx.doi.org/10.22004/ag.econ.316758`.

Acknowledgments

We thank the anonymous referees and the managing editor for their helpful comments on earlier versions of the paper. We are also grateful for the thoughtful advice given by K Hervé Dakpo. This study was funded by the Bavarian Ministry of Science and the Arts in the context of Bavarian Climate Research Network (bayklif).

4.1 Abstract

This study compares the greenhouse gas (GHG) efficiency of intensive and extensive dairy farms, and determines their GHG mitigation potential. We combine the concept of eco-efficiency with latent-class stochastic frontier analysis and the estimation of a stochastic meta-frontier. In the case of Bavaria, Germany, we find that intensive dairy farms convert GHG emissions on average more efficiently into farm economic output than their extensive counterparts. Extensive farms could, on average, reduce GHG emissions by 225 t CO_2 equivalents per year while intensive farms could reduce emissions by 130 t CO_2 equivalents without reducing their economic output.

4.2 Introduction

Agriculture accounts for about one tenth of greenhouse gas emissions in the European Union (EU) (European Statistical Office, 2018). Within the agricultural sector, dairy farming is the largest source of methane (CH_4) and nitrous oxide (N_2O) (Weiske et al., 2006). Over the past several decades, farming practices in industrial countries' dairy sector have become increasingly intensive, which is noticeable in higher stocking rates, higher feed intensity, and higher milk yields. For example, total milk production in the EU increased by about 47% from 2005 to 2016, whereas the number of milking cows remained relatively stable at around 23 million (European Commission, 2019d). The trend towards intensification in dairy farming raises the question whether intensive farms are less emission intensive than extensive dairy farms, i.e if they produce fewer emissions per unit of output. In the public and political discussion, extensive farming is often considered less harmful to nature and the environment than their intensive counterparts (Wuepper et al., 2020). However, some studies point toward opposite effects, and overall, the empirical evidence remains scarce.

Comparing emission intensity between different types of dairy systems is often done based on life cycle analysis (LCA) studies. For example, Basset-Mens et al. (2009) compare the environmental impact across four different levels of input intensities in New Zealand. Their results indicate that extensive dairy systems have fewer detrimental effects on the environment. In contrast, Capper et al. (2009) and Gerber et al. (2013) find that intensive dairy farms release fewer greenhouse gases per kg of milk produced due to higher feed conversion efficiency and animal productivity. One major limitation of LCA studies is that they are usually limited to small samples sizes due to expensive data-collection procedures. Other findings in the literature attempt to fit ecological aspects of production into environmental production models. For instance, Skevas et al. (2018d) combine farm accountancy data from a large sample of dairy farms in the Netherlands with data on nutrient surpluses to estimate the impact of farm intensification on environmental efficiency using a hyperbolic distance function. In this production model, nutrient surpluses are considered undesirable outputs, which are inherently related to the production of desirable outputs. While environmental production models are increasingly employed in the literature, they are challenged by endogeneity problems and usually rely on a set of restrictive assumptions (Atkinson et al., 2018; Kumbhakar & Tsionas, 2016), which makes them often hard to interpret and difficult to derive policy implications from. A third way to assess the environmental impact of the production activities we are following in this article is to explore the concept of eco-efficiency. Eco-efficiency describes the ability of a firm to "produce goods and services while causing minimal environmental degradation" (Kuosmanen & Kortelainen, 2005). Eco-efficiency therefore measures both economic and environmental performance of productive activities. As Kuosmanen & Kortelainen (2005) discuss, improving eco-efficiency is a cost-effective way to reduce environmental pressures and to design policies as environmental improvements are easier to implement without restricting the level of economic activity.

In this article, we combine the concept of eco-efficiency with latent-class stochastic frontier analysis to empirically measure the environmental and economic performance of extensive and intensive dairy farms. As in Dakpo et al. (2017) our main focus is on climate change-related environmental pressures (i.e., greenhouse gases). Therefore, we call our measure of interest GHG emission efficiency rather than the more general term of "eco-efficiency". By constructing a stochastic metafrontier around the identified classes, we examine the GHG mitigation potential both within and across distinct technologies.

In our case study, we focus on dairy farming in Bavaria, a federal state in southeast Germany that accounts for approximately 5 % of milk production in the EU-28 (Wimmer & Sauer, 2020). The advantage of the latent-class approach is that classifying intensive and extensive farms is purely data-driven and does not have to be made *a priori*. Thus, the approach takes into account that farm intensification is a multidimensional concept that should not be based on single indicators

(Gonzalez-Mejia et al., 2018). Our results indicate a large climate change mitigation potential for both technologies in our sample. Extensive farms could, on average, reduce GHG emissions by 225 t CO_2 equivalents per year while intensive farms could reduce emisisons by 130 t CO_2 equivalents without reducing their economic output.

The first studies concerned with eco-efficiency computed partial measures such as "economic output per unit of waste" (see Tyteca, 1996, for a review of the literature). Kuosmanen & Kortelainen (2005) argue that a more encompassing index is needed that considers different environmental pressures simultaneously. They employ data envelopment analysis (DEA) to aggregate different environmental pressures and derive a multi-dimensional eco-efficiency index for the road transportation sector in Finland. The DEA approach allows them to use objective and data-driven weights for the environmental pressures. Pérez Urdiales et al. (2016) apply the same method to the agricultural sector using cross-sectional survey data from 2010 on 50 dairy farms in Spain. Using truncated regression and bootstrapping techniques, they find that eco-efficiency is positively related to participation in training schemes and negatively related to farmer age. Employing the same data, Orea & Wall (2017) propose an SFA approach to measure eco-efficiency in order to accommodate random noise and allow for substitutability between environmental pressures. Their findings are strikingly consistent with the DEA results in Pérez Urdiales et al. (2016), confirming that SFA is an appropriate method for estimating eco-efficiency. Beltrán-Esteve et al. (2014) are the first to assess technological differences in eco-efficiency among various groups of olive producers in Spain by means of a deterministic meta-frontier approach. Further studies applying the eco-efficiency approach to the agricultural sector include Picazo-Tadeo et al. (2011), Picazo-Tadeo et al. (2012), Bonfiglio et al. (2017) and Godoy-Durán et al. (2017), all based on DEA techniques.

Our main contribution is to combine the eco-efficiency approach with the latent class stochastic frontier model. Even though the eco-efficiency concept strongly relies on the literature of productive effciency analysis, the treatment of heterogeneous technologies has gained little attention in previous studies. Exploiting the inherent link between conventional production technology and pressure-generating technology, we are able to separate distinct farm technologies (i.e., extensive vs. intensive dairy farming) as shown by the descriptive statistics of the identified groups. Further, we estimate eco-efficiency for a comprehensive data set spanning a considerably longer time horizon than most previous studies in this field. We do this by applying the GHG inventory approach based on the guidelines of the Intergovernmental Panel on Climate Change (IPCC) to recover greenhouse gas emissions from farm-level bookkeeping records. In contrast to related studies that relied on cross-sectional survey data with a limited number of observations, we are able to estimate the link between environmental pressures (i.e. emissions) and economic outcome using panel data and, hence, are able to account for unobserved heterogeneity between farms. Finally, we extend the literature on eco-efficiency by directly incorporating inefficiency determinants in the stochastic eco-efficiency frontier. While our empirical approach is based on Orea & Wall (2017), we also use a richer functional form to approximate the pressure-generating technology to allow for nonlinearities and varying returns to scale.

The remainder of this article is structured as follows. In the next section, we introduce the theoretical model that combines the eco-efficiency framework with latent-class frontier analysis. After presenting the data and the farm-level inventory approach in the subsequent section, we describe the empirical specification for estimating the pressure-generating technology. Following that, we present and discuss the results, focusing on the identification of heterogeneous technologies and the determinants of eco-efficiency. The final section concludes the study.

4.3 Conceptual framework

In this section, we first introduce the concepts of the pressure-generating technology (PGT) and eco-efficiency frontier (EEF). Further, we introduce the concept of latent-class stochastic frontier analysis within the eco-efficiency framework. Finally, we describe the stochastic meta-frontier and how it is used in our analysis.

4.3.1 A stochastic frontier approach to eco-efficiency

To assess the GHG emission efficiency of farm businesses, we rely on the eco-efficiency concept introduced by Kuosmanen & Kortelainen (2005), which compares economic success and adverse ecological impacts generated by production activity. Our goal is to measure and compare the emission-efficiency of dairy farms under different technologies. Emission-efficiency is defined as the ratio between economic output and GHG emissions, so that higher emission-efficiency implies that a farm produces more economic output with less environmental impact (Dakpo et al., 2020; Huppes & Ishikawa, 2005). The emission-efficiency scores are easy to interpret and allow us to express the GHG mitigation potential in a tangible number. A recent application of the concept of eco-efficiency to dairy farming is provided by Martinsson & Hansson (2021) using DEA. We use the pressure-generating technology to derive this measure using SFA, following Orea & Wall (2017), rather than computing the traditional eco-efficiency ratio indicators (see an overview in Tyteca, 1996). Alternative methods have been proposed in the literature to incorporate pollution in production technologies: environmentally adjusted production efficiency models, material balance principle-adjusted models, and multiple equation environmentally adjusted efficiency models (Dakpo et al., 2020; Lauwers, 2009). The latter category includes the by-production model proposed by Murty et al. (2012), which defines the global technology as the intersection of two sub-technologies – one for good outputs (e.g., milk) and one for bad outputs (e.g., greenhouse gas emissions). These models explicitly show how bad outputs are generated, and thus illustrate the mechanisms behind pollution-adjusted efficiency. Since our purpose is to measure the economic and environmental outcomes of dairy farms operated under different technologies rather than to explain how the bad outputs are generated, we use the more straightforward eco-efficiency approach in this paper. Following Picazo-Tadeo et al. (2012), we assume that we observe a set of $k = 1, \dots K$ decision-making units (DMUs), which produce economic output y each year in period $t = 1, \dots T$ by generating $n = 1, \dots N$ environmental pressures $\mathbf{s_n^t} = (s_1^1, \dots, s_N^T)$ that damage the environment. To formally describe this process, we follow Kuosmanen & Kortelainen (2005) and Kortelainen (2008) and introduce a time-dependent pressure-generating technology set ($PGTS$), which contains all economically feasible combinations of economic outputs (y) and adverse ecological impacts ($\mathbf{s}$) in period t:

$$PGTS_t = \Big[(y_t, \mathbf{s}_t) \in \mathbb{R}_+^{N+1} \Big| \text{ economic output } y_t \text{ can be generated with ecological pressures } \mathbf{s}_t \Big] \tag{4.1}$$

Given the $PGTS$ (4.1), eco-efficiency is traditionally defined as the ratio between economic output (y) and a function that aggregates environmental damages $D(\mathbf{s}_{nt})$. This quotient $y/D(.)$ is then usually reformulated as a non-parametric optimization problem, where $D(.)$ is expressed as a weighted average of $\mathbf{s}_{nt}$ and can be solved using DEA (see e.g. Bonfiglio et al., 2017; Kuosmanen & Kortelainen, 2005; Picazo-Tadeo et al., 2012, 2014). However, as Kortelainen (2008) notes, this eco-efficiency definition lacks any baseline to which to compare individual eco-efficiency levels and thus clearly deviates from the ordinary concept of relative efficiency in production economics studies. What is more, the link between environmental impact and ecological pressures is likely to be nonlinear and rather complex, which is why using a weighted average to aggregate environmental pressures, as is common practice in DEA eco-efficiency studies, is likely to be too simplistic. We therefore choose a more comprehensive parametric setting (Orea & Wall, 2017), in which we rewrite (3.1) in functional form yielding:

$$y = D(s_1,\ s_2, \ldots, s_N; \boldsymbol{\beta}) \tag{4.2}$$

where $\boldsymbol{\beta}$ represents a parameter vector to be estimated and $s_1,\ s_2, \ldots, s_N$ describe the ecological pressures that are imposed on the environment when generating economic output y. Following Orea & Wall (2017), equation (4.2) lets us assess the marginal contribution of pressure s_n to the economic outcome y of production unit k at time t. As extensively discussed in Kuosmanen & Kortelainen (2005), it is important to note that this concept relies on the presumption that economic outcome and environmental pressures are regarded from an impact-based point of view and not from a quantity-based point of view. This means that both physical inputs and physical outputs are implicitly linked to y and $\mathbf{s}$, which is why conventional productive inputs are not included in the eco-efficiency approach. Instead, "the relations between conventional inputs and desirable outputs are no longer considered explicitly, but only the ecological outcomes with respect to the economic outcomes" (Lauwers, 2009, p.1607). The absence of conventional inputs in the eco-efficiency equation implies that farms with the same amounts of outputs and GHG emissions but different levels of inputs are considered equally efficient (e.g. Tyteca, 1998). While this is an important qualification of the frontier eco-efficiency approach, it does not conflict with our goal to assess the ability of extensive and intensive dairy farms to produce milk output under minimum emissions. While the eco-efficiency approach focuses on the relationship between environmental pressures and economic outcome, the input-output technology and the pressure-generating technology are closely linked to one another, a fact we will later make use of when seeking to find underlying heterogeneous pressure-generating technologies.

So far, the EEF (3.3) does not reflect the fact that firms may not produce the maximum economically feasible level of y given their level of aggregated ecological impact $D(\mathbf{s}_{nt})$. Additionally, (4.2) is deterministic in nature and does not account for statistical noise. To overcome these shortcomings, we express (4.2) as a stochastic frontier model (Aigner et al., 1977; Meeusen & van den Broeck, 1977) and add an error term that accounts for ecological inefficiency and noise:

$$y = D(s_1,\ s_2, , \ldots , s_N; \boldsymbol{\beta})\ e^{-u+\nu} \tag{4.3}$$

where $\epsilon = -u + \nu$ represents the composite error term consisting of an ecological inefficiency component (u) and statistical noise (ν). Taking the ordinary definition of technical efficiency (TE), eco-efficiency is the ratio between observed economic output and the frontier output:

$$EE = \frac{y}{D(\boldsymbol{s}'\boldsymbol{\beta})\ e^{\nu}} = \frac{D(\boldsymbol{s}'\boldsymbol{\beta})\ e^{-u}\ e^{\nu}}{D(\boldsymbol{s}'\boldsymbol{\beta})\ e^{\nu}} = e^{-u} \tag{4.4}$$

EE can take on values between zero and one. It measures the economic output of the k-th firm relative to the maximum attainable economic outcome by a fully efficient firm producing the same degree of adverse environmental impacts. Kumbhakar et al. (2015) demonstrate that the orientation of an efficiency measure can easily be reversed. This means in our case, as we are specifically interested in the GHG mitigation rather than in the economic output expansion potential, we convert the emission efficiency (EE) measure such that it measures the emission reduction potential at a constant output level.

4.3.2 Latent-class model

The EEF in equation (4.3) implicitly assumes that firms are operated under a homogenous pressure-generating technology. Analogous to production technologies, if multiple pressure-generating technologies existed, this would result in biased parameter estimates and incorrect eco-efficiency scores (e.g., Martinez Cillero et al., 2019). It has been recognized that farms in general and dairy farms in particular are often operated under at least two distinct technologies, mostly characterized by intensive and extensive ways of production (see, e.g., Alvarez & del Corral, 2010; Orea et al., 2015; Sauer & Moreddu, 2020; Sauer & Wossink, 2013a). Since the physical relationship between inputs and outputs is closely related to the pressure-generating technology, heterogeneous technologies also affect the performance measurement with respect to the pressure-outcome combination. To allow for the presence of multiple technologies, we estimate the EEF within the latent-class stochastic frontier framework. In this framework, the pressure-generating technologies and unobserved class membership are estimated simultaneously. The identification of heterogeneous technologies in a given sample is purely data-driven but can be sharpened by the specification of separating variables, which may help explain class membership. To define the latent-class EEF, we rewrite its parametric form in (4.3) as

$$y = D(s_1, s_2, ..., s_N; \beta_{|j}) \times e^{-u_{|j}+-v_{|j}} \tag{4.5}$$

where subscript j denotes a finite number of distinct pressure-generating technologies. The number of distinct technology groups is unknown but can be tested statistically using the Akaike and Bayesian information criteria (AIC and BIC). The likelihood function for each individual farm is a weighted average of its likelihood function for each group j, weighted by the prior probability of class j membership (P_{ij}). The overall likelihood function is then given by the sum of individual likelihood functions (Greene, 2005):

$$LF_i(\theta, \delta) = \sum_{j=1}^{J} LF_{ij}(\theta_j) \times P_{ij}(\delta_j) \tag{4.6}$$

$$\log LF(\theta, \delta) = \sum_{i=1}^{N} \log \left(\sum_{j=1}^{J} LF_{ij}(\theta_j) \times P_{ij}(\delta_j) \right), \tag{4.7}$$

where θ is the vector of frontier parameters and δ_i are parameters affecting the prior probabilities. By definition, the prior probability must lie between zero and one and add up to zero, which is thus parameterized as a multinomial logit model where firm-specific variables Q_i can be used to sharpen the probability:

$$P_{ij}(\delta_j) = \frac{\exp(\delta_j' Q_i)}{\sum_j^J \exp(\delta_j' Q_i)} \tag{4.8}$$

If no separating variables are specified, Q_i is just a constant. After maximizing the overall log-likelihood function with respect to the parameters θ and δ using conventional optimization methods, the posterior probabilities of class membership can be calculated using Bayes Theorem (Greene, 2005):

$$P(j|i) = \frac{LF_{ij}(\hat{\theta}_j) \times P_{ij}(\hat{\delta}_j)}{\sum_{j=1}^{J} LF_{ij}(\hat{\theta}_j) \times P_{ij}(\hat{\delta}_j)} \tag{4.9}$$

Thus, while the prior probability only contains coefficients of the separating variables (or a constant), the posterior probability is calculated using coefficients of both separating variables and the parameters characterizing the EEF. The posterior probabilities can be interpreted as uncertainty in the class assignment of the farms. Finally, eco-efficiency is measured individually for each class j as in equation (4.4).

4.3.3 Stochastic metafrontier

Following Huang et al. (2014), we define the common underlying eco-efficiency meta-frontier for all latent classes in period t as $D_M^t(\boldsymbol{s}_{jit})$. This function is the same for all classes $j = 1, \ldots, J$. Huang et al. (2014) show that such a meta-frontier $D_M^t(\boldsymbol{s}_{jit})$ envelops all individual classes' frontiers $D_j^t(\boldsymbol{s}_{jit})$. This relation can be expressed as follows:

$$D_j^t(\boldsymbol{s}_{jit}) = D_M^t(\boldsymbol{s}_{jit})\, e^{-u_{jit}^M} \tag{4.10}$$

where $-u_{jit}^M \geq 0$ and therefore $D_M^t(.) \geq D_j^t(.)$. The ratio between the two frontiers in (4.10) is defined as the technology gap ratio (TGR):

$$TGR_{it}^j = \frac{D_j^t(\boldsymbol{s}_{jit})}{D_M^t(\boldsymbol{s}_{jit})} = e^{-u_{jit}^M} \leq 1 \tag{4.11}$$

which assesses the gap between the group frontier and meta-frontier. A TGR value of one means that a decision-making unit (DMU) i of class j has adopted the most efficient PGT_j to generate economic output at time t. A TGR value of less than one indicates a failure of the DMU to do so. In this case, the DMU could improve its efficiency by adopting the alternative technology, given its level of ecological impact. Sometimes switching to an alternative technology may not be feasible or very difficult to accomplish, i.e. adopting a new technology could imply moving to a different region. In the result section, we check if both technologies are available across all regions in the analysis. Despite potential access to all pressure-generating technologies, firms may not choose the best technology due to specific circumstances, e.g., regulations, production environments or resources (Huang et al., 2014). The pressure-outcome combination of the i^{th} firm with respect to the meta-frontier, $D_M^t(\boldsymbol{s}_{jit})$ has three components: the technology-gap ratio TGR_{it}^j, the eco-efficiency of each DMU EE_{it}^j, and the random noise component $e^{v_{jit}}$ (Huang et al., 2014), i.e.:

$$\frac{y_{jit}}{D_M^t(.)} = TGR_{it}^j \times EE_{it}^j \times e^{v_{jit}} \tag{4.12}$$

Provided the fact that a random component is obtained from the stochastic frontier estimation of the class-specific frontiers, the eco-efficiency with respect to the metaforntier efficiency (MTE) can be expressed as:

$$MTE_{jit} = \frac{y_{jit}}{D_M^t(.)\, e^{v_{jit}}} = TGR_{it}^j \times EE_{it}^j \tag{4.13}$$

In the suggested method by Huang et al. (2014), the estimation of the meta-frontier accounts for the estimation error of the class-specific eco-efficiency frontiers (4.5).

The estimation error of the class-specific frontier is then:

$$\ln \hat{D}_j^t(\boldsymbol{s}_{jit}) - \ln D_j^t(\boldsymbol{s}_{jit}) = \epsilon_{jit} - \hat{\epsilon}_{jit}. \tag{4.14}$$

Given the estimation error $v_{jit}^M = \epsilon_{jit} - \hat{\epsilon}_{jit}$, equation (4.10) can be expressed as:

$$\ln \hat{D}_j^t(\boldsymbol{s}_{jit}) = \ln D_M^t(\boldsymbol{s}_{jit}) - u_{jit}^M + v_{jit}^M \tag{4.15}$$

where $\ln \hat{D}_j^t(\boldsymbol{s}_{jit})$ are the estimates of the class-specific frontiers. This specification resembles a traditional stochastic frontier model, where the technological gap component $u_{jit}^M \geq 0$ is assumed to follow a truncated normal distribution and to be independent of v_{jit}^M. Since the class-specific frontier is supposed to be smaller than the meta-frontier $D_j^t(\boldsymbol{s}_{jit}) \leq D_M^t(\boldsymbol{s}_{jit})$, the estimated TGR must be less than or equal to 1:

$$T\hat{G}R_{it}^j = \hat{E}\left(e^{-u_{jit}^M} \mid \hat{\epsilon}_{jit}^M\right) \leq 1 \tag{4.16}$$

where $\hat{\epsilon}_{jit}^M = \ln \hat{D}_j^t(\boldsymbol{s}_{jit}) - \ln D_j^t(\boldsymbol{s}_{jit})$ are the estimated composite residuals of (4.15). Allowing for heteroskedasticity, we assume $u_{jit}^M \sim \mathcal{N}^+(\mu, \sigma^2)$ and $\mu = \omega_0 + \boldsymbol{z'_{jit}\omega}$, such that the estimated TGR is a function of the production environments z'_{jit}, for which the parameter vector $(\omega_0, \boldsymbol{\omega'})$ has to be estimated (Battese & Coelli, 1995).

4.4 Data and descriptive statistics

We use data from Bavarian dairy farms that are collected as part of the European Farm Accountancy Data Network (FADN). Bavaria is a German federal state (NUTS 1 region) located in the southeast of Germany. The data are a large unbalanced panel with 9224 observations of 1291 farms observed annually between 2005 and 2014. Farms participated on average for 7.1 years. The dataset includes farm production and economic data and is designed as a stratified sample according to farm location, size classes, and specialization of the farms. Although the data used in this study consists of a regional sample of farms, the analysis is highly relevant in a larger European context. For instance, Bavaria is one of the largest milk-producing regions within the EU (Frick & Sauer, 2018), and the Bavarian dairy farm labor structure and livestock count can be seen as representative of the bulk of European dairy farms in that average numbers are close to European averages (European Statistical Office, 2020). Hence, the results of this study might likely be informative for other regions in the EU as well.

In our analysis, we focus on climatic stresses caused by farming and include the three most important GHGs as independent variables in our model, namely CH_4,

N_2O and carbon dioxide (CO_2). GHG emissions were reconstructed using the farm-level inventory approach by Coderoni & Esposti (2018), which is largely based on macro-level international standards such as the IPCC (2006) guidelines and the United Nations Framework Convention on Climate Change (UNFCCC, 2014). The idea behind this approach is to multiply farm-level activity data (AD) by an activity-specific emission factor (EF).[1] We retrieved AD from the above-mentioned farm accountancy dataset. Regional empirical EFs were obtained from Haenel et al. (2018).[2] Following Coderoni & Esposti (2018), we set the farm-gate as system boundary, which accounts for all emissions over which the farmer has a direct control. Dairy farming in Bavaria (and other regions as well) often implies several other farming activities that might also contribute to the emission of greenhouse gases. To account for all potential sources of GHG emissions, we considered a total of eleven emission sources as described in IPCC (2006) and depicted in Table 4.3. Utilizing data from an extensive farm accountancy dataset enables us to evaluate a large sample of farms over multiple years. In contrast, most previous eco-efficiency studies that are based on LCA contain fewer sample farms (usually 50-100 max.) and cover a very limited time horizon (usually one year) (see e.g. Iribarren et al., 2011; Pérez Urdiales et al., 2016). While collecting data via LCA is very costly and thus limited to small samples, it is possible to include a rich set of farm and management specificities that might be relevant for the overall GHG emission quantity such as feed composition, housing facilities or manure management (see e.g. Zehetmeier et al., 2020). The large-scale approach chosen for this study does not allow for such detail and thus relies on a set of simplifying assumptions including homogeneous manure management and housing systems (Coderoni & Esposti, 2018), which inevitably leads to a certain degree of imprecision. The main source of heterogeneity is therefore compositional as described in Table A2 in the appendix. Hence, variations in farm-level GHG emissions arise from structural variables such as herd and crop composition. Our

[1] We follow the conventions from agricultural emissions reporting, where AD are defined in terms of areas, numbers of animals, or amounts of nitrogen (Coderoni & Esposti, 2014, 2018; Haenel et al., 2018; Njuki et al., 2016). In this context, the term AD should not be confused with farm activities in a more traditional sense such as feeding, milking or handling manure. Table A2 in the appendix summarizes the livestock and crop categories used as activity.

[2] An important aspect is that the emission factors assigned to the farm activities are constant across farms. Our approach thus assumes that the identified technology groups employ, e.g. similar manure management systems on average. If farms in one technology group systematically use less GHG-efficient manure systems than farms in the other technology group, results must be interpreted with care. Given our data, it is not possible to distinguish between such differences in this study.

analysis does not account for negative emissions through potential carbon sinks from agricultural land use (Smith et al., 2001). However, Coderoni & Esposti (2018) suggest that these negative emissions might be negligible at the farm-level compared to the other emission sources listed in Table 4.3. To compare GHG mitigation potentials of the different farms studied, CH_4 and N_2O emissions were converted to CO_2-equivalents. To that end, N_2O and CH_4 quantities were multiplied by their respective global warming potentials (298 and 34, respectively) as per the IPCC's Fifth Assessment Report (IPCC, 2013), considering the inclusion of climate carbon feedback and a 100-year time horizon. It is important to note that farm accountancy data are not primarily aimed at calculating environmental indicators such as GHG emissions and therefore the method must rely on several assumptions about farm management practices. Coderoni & Esposti (2014, 2018) present a detailed discussion about these assumptions and further technical details about the method. Although the computed GHG emissions might only be seen as a tentative approximation to the true GHG emissions, they still provide important insights regarding the overall farm-level GHG mitigation potential.[3]

Farm revenue ($yRev$), deflated to the base year 2005, enters the model as the economic outcome variable, which is defined as the value of sales (taxes included, subsidies excluded). Aggregating several farm outputs in one variable is in line with the related literature (e.g. Martinez Cillero et al., 2019; Mennig & Sauer, 2020). Since the farms in our sample are specialized dairy farms with milk revenue accounting for at least 67% of total revenue on average, we do not focus on substitutability or complementarity of individual farm outputs. Furthermore, we account for the farming environment, which is likely to influence the PGT by including regional dummy variables (r) representing five different agri-environment zones. A short description of the regions can be found in the appendix (Table 4.7). We also include a time variable (t) in our model to account for potential (ecological-)technical change.

As part of the inefficiency effects model, we include a series of variables that reflect the farmers' characteristics. These variables include the age (z_{age}) and agricultural education level of the farm manager (z_{educ}). Further, reflecting the farmers' management characteristics, we include the share of milk revenue in total revenue ($z_{smilkrev}$) as a proxy for the degree of specialization, and milk yield

[3]Their quantities (Table 4.1) are very similar to those found in previously-conducted LCAs such as, e.g., Basset-Mens et al. (2009), O'Brien et al. (2012) and Zehetmeier et al. (2014).

($z_{milkyield}$) in the inefficiency effects model. Finally, we expect that participating in agri-environment schemes (z_{aes}) affects farms' emission inefficiency.

As for the separating variables of the PGT, we make use of the close link between production technology and pressure-generating technology and presume that classical production-related separating variables often found in the literature are equally capable of separating different PGTs. Hence, we follow Alvarez & del Corral (2010) and Sauer & Wossink (2013a) and include the stocking rate (number of cows per hectare land, $qCowHA$) and feeding intensity (purchased concentrated feed per cow, $qFeedHA$) as separating variables into our model.

Table 4.1: Descriptive statistics (N = 2473).

Variable	**Description**	**Mean**	**SD**	**Min.**	**Max.**
$yRev$	Farm revenue (€1,000)	154.58	87.29	11.09	1139.32
CH_4	Methane emissions (t CO_{2e})	417.79	202.4	62.7	2421.88
N_2O	Nitrous oxide emissions (t CO_{2e})	87.03	47.45	10.59	471.37
CO_2	Carbon dioxide emissions (t CO_{2e})	28.85	19.89	1.83	310.41
t	Time (years)	7.14	2.48	3	10
rA	Region A (1 if region A, 0 otherwise)	0.63	0.48	0	1
rB	Region B (1 if region B, 0 otherwise)	0.15	0.36	0	1
rC	Region C (1 if region C, 0 otherwise)	0.11	0.31	0	1
rD	Region D (1 if region D, 0 otherwise)	0.01	0.08	0	1
rE	Region E (1 if region E, 0 otherwise)	0.11	0.31	0	1
Separating Variables					
$qCowHA$	Stocking rate (no. dairy cows/ha)	1.48	0.52	0.2	6
$qFeedInt$	Feeding intensity (purchased feed in €/cow)	284.17	137.55	1.84	2303.79
Efficiency Determinants					
$z_{milkyield}$	Milk yield (liters milk per cow)	6759.47	1108.89	2089.23	10494.65
$z_{sMilkRev}$	Revenue share of milk sales	0.78	0.06	0.66	0.99
z_{age}	Age of the farm manager (years)	49.13	9.3	19	85
z_{educ}	Agricultural education (1 = Advanced technician/University degree, 0 otherwise)	0.3	0.46	0	1
z_{aes}	Agri-environmental payment (1 = receives payments, 0 otherwise)	0.63	0.48	0	1

4.5 Empirical specification

The empirical version of the latent-class eco-efficiency frontier (4.5) takes on the second-order flexible translog functional form:

$$\begin{aligned}\ln yRev_{it} =& \beta_0|_j + \sum_{l=1}^{3} \beta_l|_j \ \ln s_{lit} + \frac{1}{2}\sum_{l=1}^{3}\sum_{l=1}^{K} \beta_{lk}|_j \ \ln s_{lit} \ln s_{kit} + \sum_{z=1}^{4} \phi_z|_j \ r_{zi} \\ &+ \beta_t|_j \ t + \frac{1}{2}\beta_{tt}|_j \ t^2 + \sum_{l=1}^{3} \beta_{lt}|_j \ \ln s_{lit} \ t + \nu_{it}|_j - u_{it}|_j \end{aligned} \tag{4.17}$$

where r_{zi} are the regional dummies and t is a time trend. The β's are unknown coefficients to be estimated. Emissions $\boldsymbol{s} \in \{CH_4, N_2O, CO_2\}$ have been mean-scaled prior to estimation, which allows the first-order coefficients to be interpreted as emission elasticities for a farm characterized by an emission portfolio equal to the sample arithmetic mean. As for the stochastic part of the model, ν corresponds to normally distributed stochastic noise, u is the inefficiency term following a positive truncated normal distribution with constant scale parameter σ_u^2 and a location parameter μ, which depends on the set of inefficiency determinants z following Battese & Coelli (1995).

Regarding the specification of the meta-frontier model, we assume a translog functional form as stated earlier:

$$\begin{aligned}\ln \widehat{D(\boldsymbol{s'\beta}|_{\boldsymbol{j}})} =& \beta_0^M + \sum_{l=1}^{L} \beta_l^M \ \ln s_{lit} + \frac{1}{2}\sum_{l=1}^{L}\sum_{l=1}^{K} \beta_{lk}^M \ \ln s_{lit} \ln s_{kit} + \\ & \beta_t^M \ t + \frac{1}{2}\beta_{tt}^M \ t^2 + \sum_{l=1}^{L} \beta_{lt}^M \ \ln s_{lit} \ t + \nu_{it}^M - u_{it}^M \end{aligned} \tag{4.18}$$

where $\ln \widehat{D(\boldsymbol{s'\beta}|_{\boldsymbol{j}})}$ corresponds to the predicted revenues on their respective class eco-efficiency frontier. Location dummies are omitted from the meta-frontier, which is the same at all regions, while all assumptions from the previous section apply equally to equation (4.18). Instead, we include regional dummies representing the production environment of farms in the TGR term of the stochastic meta-frontier. This is because the technology gap represented by the one-sided u is due to the choice of a particular technology that depends on the production environment (Huang et al., 2014).

As defined above, emission-efficiency measures the economic output of a firm relative to the maximum attainable economic outcome given the same level of adverse environmental impacts. For policy implications, it may be more meaningful to measure the extent to which GHG emissions could be decreased without changing the economic output. Kumbhakar et al. (2015) demonstrate that the output-oriented approach in Eq. 3.4 can be reformulated as an *'emission-oriented'* approach: $y = D(\mathbf{s}e^{-\eta}, \boldsymbol{\beta})e^{\nu}$, where η is the efficiency measure with respect to emission reduction (i.e. GHG mitigation potential). They then show that the output-oriented efficiency measure (u) for a translog function can be expressed in terms of η in the following way:

$$u = \eta \left[\sum_j \beta_j + \sum_j \beta_{jt} t + \sum_j \left(\sum_k \beta_{jk} \ln x_k \right) \right] + \frac{1}{2}\eta^2 \sum_j \sum_k \beta_{jk} \tag{4.19}$$

Solving the quadratic equation (4.19) for η yields the desired "emission-oriented" efficiency measure. In the following results section, we will explicitly focus on efficiency scores that express the GHG mitigation potential (η). Outcome-oriented efficiency scores (u) can be obtained from the authors upon request. We note that the signs and significances of the inefficiency determinants can be interpreted using both perspectives: A variable that is positively related to emission-efficiency is also positively related to GHG mitigation.

4.6 Empirical results and discussion

In this section, we first present our findings regarding the detection of distinct intensity classes and discuss implications. We then depict and discuss class-specific emission efficiency scores as well as GHG mitigation potentials. Finally, we turn to the results from the meta-frontier estimation.

4.6.1 Heterogeneous technologies and class separation

We used the software LIMDEP 10 to implement the estimation of the latent-class eco-efficiency frontier (4.5), where time-varying and farm-specific inefficiency is expressed as $u_{it}|j = g(z_{it})|j \times |U_i|$. This specification ensures that farm-specific inefficiency is not independent over time.[4] As a robustness check, we also estimated the latent-class frontier model specified in Greene (2005), which assumes independence of the efficiency term over time. Overall, 74% of farm-observations are classified into the same technology groups as in the above mentioned specification.

The number of distinct technology classes is unknown and must be specified prior to the estimation. To determine which group number best represents the data, we compute AIC and BIC statistics (e.g., Alvarez & del Corral, 2010; Orea & Kumbhakar, 2004). The estimation with four and three groups specified did not converge, which is likely due to over-specification of the model (e.g., Orea & Kumbhakar, 2004). The AIC and BIC values for the model with two groups (-11,939 and -11,532) are both lower than the ones for the pooled model (-11,477 and -11,284), indicating that the data supports the existence of two distinct technology

[4] However, this approach does not account for adjustment costs, which arise from adapting the production processes (Stefanou, 2009). In stochastic frontier analysis, dynamic efficiency can be evaluated using autocorrelated ineffiency terms (Ahn & Sickles, 2000; Emvalomatis et al., 2011; Tsionas, 2006). Recent applications in dairy farming are provided by Skevas et al. (2018a,b,c). Alternatively, the dynamic structure of the decision making process can also be modeled by including gross investments in the frontier (e.g. Minviel & Sipiläinen, 2018; Serra et al., 2011).

groups. For comparison reasons, we report the parameter estimates for both the pooled (column 1) and the latent-class models (columns 2 and 3) in the appendix (Table 4.8.3). Since all variables have been divided by their sample means prior to estimation, the first-order coefficients can be interpreted as emission elasticities evaluated at the sample mean. This is because the log of demeaned variables takes on a value of 0 at the sample mean, and thus observation-specific values drop out in the first derivatives of the production frontier. Note that the first-order parameters represent elasticities at the sample mean of all farm observations, regardless of group membership.

Consistent with expectations, the elasticities are significantly positive across all models. For example, a one-percent increase in CH_4 release is associated with a 0.73% increase in revenue in technology class I and with a 0.79% increase in revenue in technology class II. The mean elasticities are largely consistent between the pooled model and the two identified classes. Technological differences are particularly visible in the constant terms as well as in the second-order parameters of the eco-efficiency frontiers (Appendix 4.8.3). For example, the marginal effect of CH_4 increases in CO_2 emissions in class II, whereas CH_4 affects economic output independent of CO_2-levels in class I. Moreover, eco-technical progress indicated by the pooled model seems to be driven by class I, as both the linear time variable and its squared term are statistically insignificant in class II. Finally, regional dummies are statistically significant in class I but insignificant in class II. Since we use the most favorable agricultural region in terms of farming conditions (*"Gäu"*) as a baseline region in the estimation, the negative coefficients in class I can be expected. The insignificant parameters in class II imply that local conditions are less important for farms operated under this technology.

The prior probabilities, which are calculated using the coefficients of the separating variables, are 52% for technology class I and 48% for class II on average, indicating that the group sizes are roughly balanced and that the assumed technology specification is relevant. The posterior probabilities of belonging to class I or class II are computed using the coefficients of both separating variables and the parameters of the corresponding eco-efficiency frontier. Here, we report the average posterior probabilities of belonging to a particular class for those farms being assigned to the same class: farms that are assigned to class I (class II) have an 85% (88%) posterior probability of belonging to class I (class II). These posterior probabilities indicate that the uncertainty in the class assignment is relatively low. Sauer &

Wossink (2013a) and Orea et al. (2015) find very similar posterior probabilities in a latent-class production frontier applied to dairy farms.

Next, we examine how farms classified to distinct technology classes can be characterized. The separating variables are statistically significant and negative. This result implies that both stocking rate and the feed intensity decrease the likelihood of a farm being assigned to class I. Thus, we can expect that class I represents an extensive technology while class II represents a more intensive technology. To confirm this, we report class-specific descriptive statistics in Table 4.2. Corresponding to the prior probabilities, the number of farms assigned to each group is quite similar, with 5,062 farm observations being assigned to class I and 4,162 farm observations being assigned to class II. Farms in class I are characterized by smaller amounts of gross outputs, both in absolute terms and relative to the number of cows and land operated. Additionally, farms in this group have on average smaller herd sizes, more hectares of land use, lower stocking rates, lower feed purchases per cow, and lower capital usage per cow. The third column of Table 4.2 indicates that the differences are all statistically significant at the 1% significance level, according to Student's t-tests. Thus, the descriptive statistics confirm that class I represents the extensive technology group and class II the intensive one.

The result that the latent technologies reflect different production intensities is in line with, e.g., Alvarez et al. (2012) or Martinez Cillero et al. (2018), who estimate latent-class production frontiers rather than pressure-generating technologies. Although it is not possible to connect these results directly to differences in specific farming procedures and management practices concerning, e.g. manure management technologies, cropping systems, machinery, milking technology, feed management, veterinary services, etc., the latent-class approach enables us to compare greenhouse gas efficiency between the distinct technology groups.

Clearly, the latent classes that were found based on the PGT reflect two groups that show considerable differences with respect to their production intensity, as depicted in Table 4.2. This finding is in line with previous research that finds a positive correlation between input use and environmental pressures, such as nutrient runoff or GHG emissions (Guerci et al., 2013; Lötjönen et al., 2020; Orea & Wall, 2017; Pérez Urdiales et al., 2016).

Table 4.2: Summary statistics and differences of mean farm characteristics by technology class.

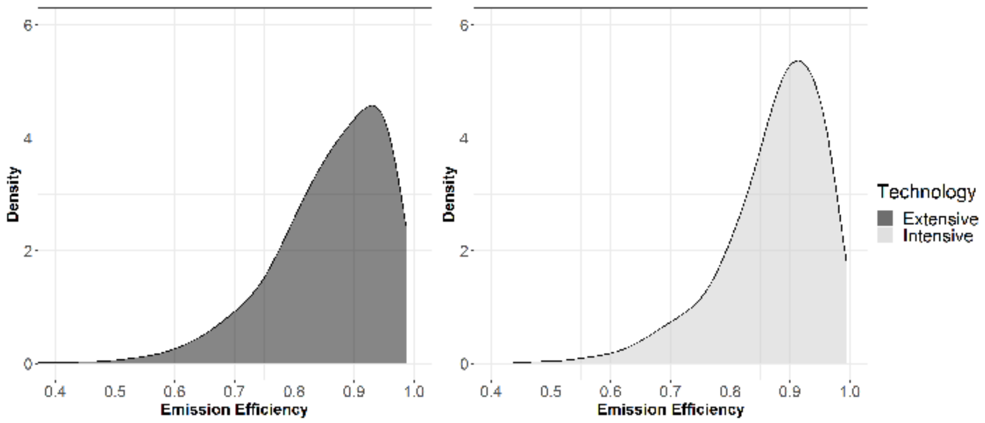
	Class I	**Class II**	**Difference**
Observations	5,062	4,162	
Total farm gross output (€)	41,406.72	56,221.76	14815.038***
GHG emissions (tons CO_{2eq})	507.121	622.461	115.34***
Cows (units)	58	73	15***
Land (ha)	49.064	46.871	-2.193***
Total farm gross output per cow (€)	712.148	770.67	58.522***
Total farm gross output per hectare (€)	911.009	1,299.869	388.86***
Stocking rate (Cows/ha)	1.306	1.701	0.395***
Milk (l)	257,338.2	368,329.8	110991.64***
Milk per cow (l)	6,310.643	7,305.355	994.712***
Milk per hectare (l)	5,612.597	8,532.815	2920.218***
Concentrate feed per cow (€)	229.301	350.896	121.595***
Capital per cow (€)	448.638	471.934	23.296***
Class name	Extensive	Intensive	

Single, double, and triple asterisks (*, **, ***) indicate [statistical] significance at the 10%, 5%, and 1% level.

4.6.2 GHG emission efficiency by production intensity

As shown in the previous section, we are able to identify two intensity classes that demonstrate distinctively different GHG emission generating technologies. Furthermore, we assumed that farms may be inefficient in that they could reduce their emissions while keeping their revenue constant. Figure 4.1 summarizes the groupwise GHG emission efficiency estimates.

In terms of the relative efficiency scores, we see that both intensive and extensive technology show similar results, in that their efficiency scores are similarly distributed with a mean of 87% and 86% respectively. However, given the fact that intensive farms emit on average more greenhouse gases than extensive farms (see Table 4.2), EE scores translate into different absolute GHG mitigation potentials. If extensive dairy farms were fully emission efficient, they could reduce their GHG emissions, on average, by 67.7 tons CO_2-equivalents (CO_{2eq}) per year while their intensive counterparts could mitigate 77.2 tons CO_{2eq}. Considering the case in which all 1291 dairy farms in our sample had been fully emission efficient in terms of their respective class frontiers, a total of 664,142 tons CO_{2eq} could have been omitted without compromising economic output in the period considered from 2005 to 2014.



	Extensive Class	Intensive Class
Emission efficiency w.r.t. the class frontiers		
Mean	0.8614	0.872
Std. dev.	0.0958	0.0863
5% quantile	0.6792	0.6991
95% quantile	0.9736	0.9738
GHG mitigation potential (t CO_{2eq})		
Mean	67.71	77.22
Std. dev.	56.37	65.24
5% quantile	10.92	12.83
95% quantile	177.71	202.73

Figure 4.1: Emission efficiency and GHG mitigation potential with respect to the class frontiers.

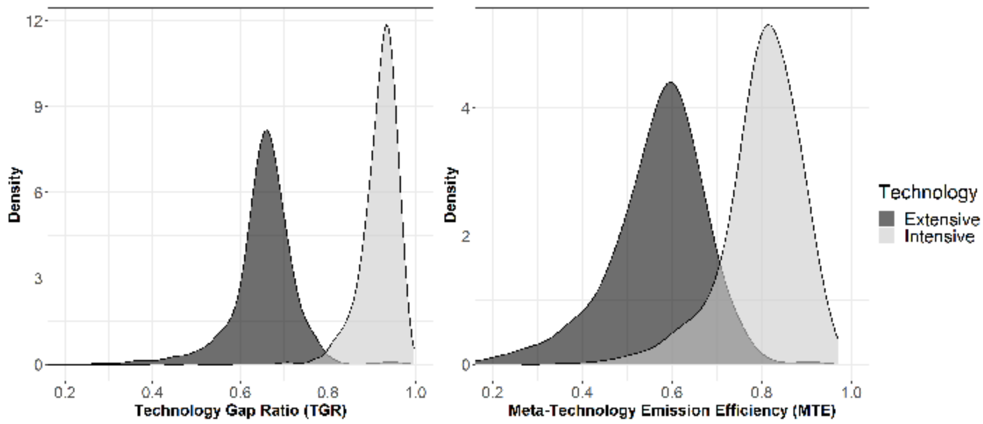
Furthermore, our empirical framework allows us to evaluate potential GHG emission inefficiency drivers. As outlined in the data section, we included five farm management-related variables into the latent-class model. Estimation results can be found in Appendix 4.8.3. We find that farmer age and education do not affect emission efficiency. However, the degree of specialization expressed as the share of milk sales on total revenue is positively related to EE for both groups. However, the association is considerably stronger for intensive farms. A similar picture is found with respect to milk yield. It is positively associated with EE for both groups, while the association is stronger for intensive farms. Overall, the effect size seems to be largest for milk yield. Finally, participation in agri-environment programs appears to have a negative effect on emission efficiency for both the intensive and extensive classes.

As we have hypothesized, farms can indeed maintain their level of production while simultaneously reducing their negative impact on the climate, i.e., reducing their GHG emissions, which holds true for both technology classes. Although both groups show a similar level of emission efficiency in relative terms, the fact that intensive farms emit on average more GHGs leads to a considerably higher mitigation potential for this group with respect to their own PGT. Previous studies on eco-efficiency have also found considerable eco-inefficiency across different farming systems such as olive growing (Beltrán-Esteve et al., 2014; Picazo-Tadeo et al., 2012), rain-fed agriculture (Gadanakis et al., 2015; Picazo-Tadeo et al., 2011), horticulture (Godoy-Durán et al., 2017), and dairy farming (Iribarren et al., 2011; Orea & Wall, 2017; Pérez Urdiales et al., 2016; Shortall & Barnes, 2013). In their study, Shortall & Barnes (2013) explicitly focus on greenhouse emissions from dairy farms in Scotland and find similar results. Milk yield appears to positively affect emission efficiency, while the education level does not appear to have an effect. In their analysis, however, they find that participation in AES is not or only very weakly associated with EE, while we find a clear negative association between participation in agri-environment schemes (AES) and EE. This might be due to two reasons. First, most agri-environment schemes in Bavaria aim at the extensification of production (ART, 2019), which probably decreases both the production level and the environmental pressures. If economic outputs decrease over-proportionally compared to emissions, emission efficiency might be negatively affected as the ratio of emissions over economic output decreases. Second, most AES are not explicitly aimed at reducing greenhouse gas emissions (ART, 2019; Stetter et al., 2022a). These measures might improve the ecological performance of farms with respect to other environmental indicators such as biodiversity and water quality, but might worsen their climatic impact. There is likely no one-fits-all policy solution regarding the overall environmental impact of farms, which in turn, means a trade-off between different environmental goals, which might be solved by prioritizing certain goals over others and by being aware of potential ramifications this practice could have.

Agri-environmental schemes in our study area do not appear to promote emission efficiency among dairy farms across the intensity classes. The highest positive association with EE has the milk yield in both groups, which can be regarded as a proxy for managerial capability. A higher milk yield indicates not only an increased productivity level but also an effective way to reduce greenhouse gas emissions (especially CH_4) because fewer cows are needed for the same level of milk production (Shortall & Barnes, 2013). In order to achieve the EU policy goal of tackling climate change (European Commission, 2019b,c) another way forward could be to curb production and thus reduce GHG emissions. However, limiting the production of agricultural goods is clearly counter the objective of meeting the globally increasing demand for agricultural products (European Commission, 2019c). Such a strategy is unlikely to reconcile the two policy goals of mitigating climate change and keeping up agricultural production levels. Instead, our results suggest that there is large potential for climate change mitigation without risking the economic viability of farms. The increased emission efficiency of intensive dairy farms implies that supporting a sustainable intensification could help to reach both economic and climate goals. In this context, it seems essential to improve the effectiveness of agri-environmental schemes, since they are negatively related to farms' eco-efficiency in our empirical application. The managerial capacity of farmers, on the other hand, is associated with an increase in emission efficiency. Therefore, e.g., advisory and training programs could be interesting policy instruments to improve farmers' managerial capability (Picazo-Tadeo et al., 2014). The ability to produce farm sales with minimum environmental impact such as GHG emissions is also influenced by TE (i.e. the ability to maximize output at given input levels) of the farms as shown by, e.g., Picazo-Tadeo et al. (2011) and Shortall & Barnes (2013). Furthermore, previous research has shown that intensive farms are more technical efficient than extensive farms (Alvarez et al., 2012). Hence, our results suggest that this difference might also translate to benefits in terms of GHG-efficiency.

Note that the results and conclusions should be interpreted within the European dairy sector with relatively small dairy farms compared to, e.g. the U.S. or Canada, where dairies with 1,000 to 5,000 cows are not uncommon. Such production systems are operated under very distinct conditions, e.g. feedlot operations. In the U.S./Canadian context, Le et al. (2020) find that TE and emission-adjusted TE are very high across dairy farms in Ontario. They conclude that minimizing GHG emissions may not be inconsistent with maximizing output for given levels of input. Njuki et al. (2016) also account for the emission of GHGs in their study on Northwestern U.S. dairy farms. Although not explicitly accounting for latent production technologies, they look at different size classes and find that large dairy operations are environmentally more inefficient regarding GHG emissions compared to their smaller counterparts, which is somewhat contrary to our findings. Overall, empirical work on the GHG emission efficiency in the U.S./Canadian dairy farming context remain scarce.

4.6.3 Meta-technology frontier and GHG mitigation potential



	Extensive Class	Intensive Class
Technology gap ratio		
Mean	0.5931	0.9013
Std. dev.	0.0753	0.0509
5% quantile	0.4639	0.8042
95% quantile	0.6994	0.9625
Emission efficiency w.r.t. the meta-frontier		
Mean	0.5157	0.7889
Std. dev.	0.1056	0.0875
5% quantile	0.3195	0.622
95% quantile	0.6682	0.9099
GHG mitigation potential (t CO_{2eq}) w.r.t. meta-frontier		
Mean	217.8	127.83
Std. dev.	113.75	89.54
5% quantile	78.33	30.93
95% quantile	428.55	292.5

Figure 4.2: Technology gap ratio, emission efficiency and the GHG mitigation potential in tons CO_2-equivalents with respect to the meta-frontier.

We have found heterogeneous pressure-generating technologies among the observed dairy farms and different GHG mitigation potentials across farms and technologies. Next, we seek to analyze the farms' emission performance with respect to the meta-frontier, thus allowing to compare the two technology classes directly. Appendix Table 4.6 shows the summary statistics of the stochastic meta-frontier estimation of equation (4.18). As expected, the SMF estimates reveal significant environmental impact of the location on the meta-frontier production function. Namely, farms not located in the region d operate under inferior emission-generating technology compared to the reference, i.e., they are further away from the meta-frontier. As highlighted in the left panel of Figure 4.2, intensive farms seem to be more efficient in adopting the best available emission-generating technology (i.e., they are closer to the meta-technology), in that their average TGR lies at 0.9 compared to 0.6 for extensive farms. Hence, if all intensive farms were operated under the meta-technology, they could obtain the same revenue while producing on average 90% of the emissions they generate using the best practices available to the intensive class of farms. In other words, only a few farms operating under the intensive technologies would obtain higher efficiency levels by switching to the extensive technology. Given this result, we can conclude that the intensive farming system has the more emission-efficient technology. Because the meta-technology envelops both technologies, the intensive class it not *a priori* closer to the meta-frontier than the extensive class at all points. Indeed, the left panel of Figure 4.2 shows an overlapping area between intensive and extensive farms. However, provided the small size of this area, where the extensive technology might be superior to the intensive technology, it does not change the general interpretation of the model.

The MTE, i.e., the product of the TGR and class-wise emission efficiency describes an individual farm's distance to the meta-frontier. Figure 4.2 compares the meta-technology emission efficiency of the two groups. As hypothesized, extensive farms are overall less emission-efficient regarding the meta-frontier than intensive farms, i.e., generate more GHG emissions per economic output than their intensive counterparts. The mean MTE score of 51.5% for extensive farms signifies that they could obtain the same production level while generating only 51.5% of their current GHG emissions if they were fully emission-efficient with respect to the meta-frontier. The mean MTE score for the intensive farm group is with 78.9% less pronounced.

Expressing the meta-frontier inefficiency–and hence the GHG mitigation potential without minimizing farm revenues–in absolute terms, extensive farms could mitigate on average 225 t CO_{2eq} per year and intensive farms could mitigate on average 130 t CO_{2eq}per year. Summing up these GHG mitigation potentials for our sample of approximately thousand farms per year over ten years reveals the potential to mitigate approximately 1.7 mil. t CO_{2eq}. Of course, this calculation assumes that all farms have access to the available meta-technology such that they could potentially adopt the superior technology (in terms of emission efficiency). One indicator for this assumption to hold true could be the presence (and thus availability) of both intensive and extensive technologies in all studied subregions. Appendix Table 4.7 shows the prevalence of each technology in the respective regions and can be seen as indicative of the fact that all farms in the sample might have indeed access to both technologies.

Looking at the GHG mitigation potential with respect to an overall potential across technologies rather than within each distinct technology, our findings show that the intensive class is closer to the emission-generating meta-technology frontier than extensive farms. Hence, if we change the point of reference (from class frontier to meta-frontier), intensive farms become favorable with respect to emission efficiency. This finding is in line with previous research (Crosson et al., 2011; Gerber et al., 2013; Guerci et al., 2013). Currently, large parts of the general public and legislators regard extensively managed, small farms as desirable from a sustainability perspective (Wuepper et al., 2020), which is why the EU actively promotes the extensification of farm businesses within its Common Agricultural (CAP) (European Commission, 2013). In light of our findings, the concept of rewarding extensification should be critically assessed with respect to climate change mitigation. Switching technologies from extensive to intensive dairy farming might be one way to improve the economic-ecological performance of farms, implying a sustainable intensification of the agricultural sector. However, we emphasize that this result refers to GHG emissions only and legislators should be careful in promoting such a step because this might have negative consequences for other important environmental aspects of farming such as biodiversity, animal welfare or soil health. For instance, Norris (2008) note that intensification of agriculture has been a major driver of biodiversity loss. Nonetheless, there could also be synergies across objectives, e.g., decreased levels of nitrogen runoff are also likely to cause decreased greenhouse gas emissions (Haenel et al., 2018).

4.7 Conclusion

In this article, we evaluated the emission efficiency of distinct technologies in dairy farming. Our approach is based on a pressure-generating technology that describes the relationship between agricultural revenue and released greenhouse gases (GHG) (or any other environmental pressure) at the farm level. Thus, emission efficiency measures the ability of farms to generate revenue while causing minimal GHG emissions. We estimated a eco-efficiency frontier in a latent-class stochastic frontier framework to identify unobserved heterogeneity in the pressure-generating technology. Our results show that dairy farms in our sample can be classified into two distinct technology classes. The two classes can mainly be distinguished by the input intensity of the respective farms in each class. The extensive and the intensive farms show very similar emission efficiency scores when evaluated against their class-specific frontier. The meta-frontier envelops both identified technologies and reveals that extensive farms are overall less emission-efficient than intensive farms: without losses in economic output, extensive farms could reduce their GHG emissions to 51.5% of current levels when choosing the most efficient technology, compared to 78.9% for intensive farms. Overall, up to 1.7 mio. t CO_{2eq} could have been saved in our sample between 2005 and 2014 without reducing economic outcome. Overall, our findings show that technology differences matter, not only with respect to TE, as suggested by previous research (Alvarez et al., 2012; Martinez Cillero et al., 2018), but also with respect to emission-efficiency. This fact has been largely neglected in previous research on environmental and eco-efficiency.

Naturally, dairy farms produce other environmental pressures in addition to greenhouse gases, such as nutrient surpluses, biodiversity losses or pesticide risks. One limitation of our study is that we cannot quantify trade-offs and interaction effects across different environmental pressures (marginal change of pressure substitution). Furthermore, GHG emissions are only considered within the farm-gate. Thus, we cannot account for indirect emissions from farming through upstream activities such as fertilizer production, although they might also play an important role with respect to global warming. In addition, the choice of extensive vs. intensive agricultural production has further implications on GHG emissions through its impact on the global food system. For example, on the global scale, the higher demand for maize and concentrate feed may be associated with land use changes, which would increase GHG emissions following intensification (Styles et al., 2018). Along these lines, following Kuosmanen & Kortelainen (2005), one major critique of the suggested approach is that even if the relative level of climatic pressure is low relative to economic output, the absolute climatic pressure (i.e., GHG emissions) can still exceed the carrying capacity of the atmosphere.

Our study showed that the eco-efficiency approach can be employed in a latent class framework to account for production heterogeneity in environmental efficiency models. Combining the latent class approach with the by-production model by Murty et al. (2012) would by a highly valuable extension of the current analysis, to provide further insights into the mechanisms how bad outputs are generated in distinct farm technologies. It would further allow to estimate separate efficiency scores for each good output, bad output, and input (see, e.g., Ait Sidhoum et al., 2019, for a recent application in agricultural economics). The active development of by-production models in the stochastic frontier setting (Lai & Kumbhakar, 2021) is a promising way forward, but to date, stochastic by-production models have not been extended to a latent-class setting to distinguish between different unobserved technology classes. Moreover, it is important to highlight that our approach measures emission efficiency in a static framework, which does not take into account the time interdependence of production decisions. Reducing greenhouse gas emissions may require changes in quasi-fixed inputs and reallocation of variable inputs, which involves adjustment costs. As the adjustment costs may differ between extensive and intensive farms, such a comparison would be a valuable extension to the present study.

What is more, policy conclusions could be substantiated if more detailed data on specific farming procedures and management practices were added to the European farm accountancy data network. Also, the analysis could be extended to additional environmental pressures, allowing for the quantification of the abovementioned trade-offs and interactions thus a more holistic assessment of the economic-ecological performance of farms. Finally, it would be interesting if the analysis were replicated for other regions and different farm types, such as other livestock farms or crop farms, which would provide additional insights into the discussion on extensive and intensive farming practices.

4.8 Appendices

4.8.1 Greenhouse gas sources, activity data and emission factors for the computation of farm-level GHG emissions

Table 4.3: Summary of greenhouse gas sources, activity data and utilized emission factors for the computation of farm-level GHG emissions. A detailed description of the methodology can be found in Haenel et al. (2018).

Gas	Emission source	Activity data		EF
		FAD	Other	
	Livestock			
CH_4	Enteric fermentation	Livestock count[a]		regional
CH_4	Manure management	Livestock count[a]		regional
N_2O	Manure management (direct and indirect)	Livestock count[a]	NH_3 and NO emission factors (indirect)	regional
	Crop cultivation			
N_2O	Use of synthetic fertilizers	Fertilizer expenditures	State-level shares, prices	regional
N_2O	Use of organic fertilizers	Livestock count[a]	N excretion factors	regional
N_2O	Atmospheric deposition of reactive nitrogen	Livestock count & fertilizer expenditures	$Frac_{Gas}$	default & regional
N_2O	Leaching and surface runoff	Livestock count[a] & fertilizer expenditures	$Frac_{Leach}$	default & regional
N_2O	Crop residues	Crop area and yield[a]	Various constants	default
CO_2	Urea application	Fertilizer expenditures	State-level shares, prices	regional
CO_2	Liming	Fertilizer expenditures	State-level shares, prices	regional
	Energy use			
CO_2	Fuel combustion	Fuel expenditures	Diesel price	default

[a] See Table 4.4 for a more detailed description of the farm-level data used for the estimation of the GHG emissions.

Table 4.4: Description of the livestock and crop categories used as activity data for the approximation of the GHG emissions. Each sub-category gets assigned separate emission factors.

Emission Source/ Activity Data	
Livestock The following distinction has been made with respect to nitrogen excretion rates as well as enteric fermentation rates at the farm level.	
a) Cattle:	– Calves – Male beef cattle – Heifers – Mature males > 2 years – Dairy cows – Suckler cows
b) Pigs:	– Weaners – Fattening pigs – Boars – Sows
c) Poultry:	– Laying hens – Pullets – Broilers – Other poultry
d) Sheep	
e) Horses:	– Heavy horses – Ponies – Light horses
Agricultural crops The following distinction has been made with respect to N_2O and N_2 emissions stemming from crops at the farm level. This implies yields and farmed area.	
	– Winter wheat – Spring wheat – Rye – Winter barley – Spring barley – Oat – Triticale – Grain maize – Maize for silage – Rape – Sugar beet – Potatoes – Grass (fodder production) – Clover, grass clover leys, clover alfalfa mixtures – Meadows – Pastures

4.8.2 Description of the agricultural regions

Region	**Name**	**Description**
Region A	(Pre-)Alpine Land & Eastern Low Mountain Range	Mountainous terrain, mostly grassland-based, region at altitude, relatively low yield potential, medium to high precipitation
Region B	North Bavarian Hill Lands	Low precipitation, medium to bad soil conditions
Region C	Tertiary Terrain	Favorable production conditions, medium to high precipitation, relatively high yield potential.
Region D	Gaeugebiete	Very good soils, favorable weather conditions, very high yield potential.
Region E	Jura Mountains and Franconian Plateau	Dry region with very low precipitation, low yield potential.

4.8.3 Results of the pooled as well as of the latent-class stochastic frontier estimation separated by classes.

Table 4.5: Results of the pooled as well as of the latent-class stochastic frontier estimation separated by classes.

	Pooled		Class I		Class II	
Variable	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.
Pressure Conversion Frontier						
$Constant$	0.378***	0.020	0.428***	0.048	0.199***	0.037
$\ln CH_4$	0.791***	0.014	0.727***	0.027	0.787***	0.020
$\ln N_2O$	0.069***	0.014	0.099***	0.027	0.044**	0.022
$\ln CO_2$	0.165***	0.006	0.178***	0.011	0.157***	0.009
$.5 \times \ln CH_4{}^2$	0.114	0.071	0.276*	0.159	-0.030	0.098
$.5 \times \ln N_2O{}^2$	0.639***	0.113	0.403**	0.169	0.988***	0.199
$.5 \times \ln CO_2{}^2$	0.162***	0.017	0.144***	0.027	0.192***	0.027
$\ln CH_4 \times \ln N_2O$	-0.367***	0.089	-0.351**	0.153	-0.443***	0.145
$\ln CH_4 \times \ln CO_2$	0.114***	0.035	-0.025	0.053	0.301***	0.058
$\ln N_2O \times \ln CO_2$	-0.220***	0.040	-0.050	0.063	-0.448***	0.068
t	1.40E-04	0.001	-7.37E-05	0.001	4.10E-04	0.001
$.5 \times t^2$	0.001***	4.40E-04	0.001**	0.001	0.001	0.001
$\ln CH_4 \times t$	-0.017***	0.004	-0.011**	0.005	-0.027***	0.007
$\ln N2_O \times t$	0.021***	0.004	0.011*	0.006	0.034***	0.007
$\ln CO_2 \times t$	-0.005***	0.002	-0.003	0.003	-0.006**	0.003
rA	-0.212***	0.020	-0.401***	0.049	-0.008	0.037
rB	-0.216***	0.021	-0.367***	0.049	0.004	0.039
rC	-0.191***	0.021	-0.349***	0.051	0.003	0.038
rE	-0.196***	0.021	-0.375***	0.049	0.043	0.039
Inefficiency determinants						
$Constant$	15.002	1.93E+05	15.371	3.08E+05	27.836	3.13E+06
$z_{milkyield}$	-1.664***	0.032	-1.799***	0.056	-3.058***	0.157
z_{age}	0.002	0.001	0.003	0.002	4.80E-04	0.003
z_{educ}	-0.008	0.042	-0.114	0.090	0.075	0.072
z_{aes}	0.131***	0.018	0.159***	0.036	0.181***	0.047
$z_{sMilkRev}$	-0.708***	0.100	-0.290*	0.160	-2.121***	0.372
Separating variables						
$Constant$			6.874***	0.833	-	
$qCowHA$			-1.893***	0.322	-	
$qFeedInt$			-0.014***	0.002	-	
Class probabilities						
Prior probability			0.523		0.477	
Posterior probability			0.851		0.879	
Model diagnostics						
AIC	-11,477			-11,939		
BIC	-11,284			-11,532		

Single, double, and triple asterisks (*, **, ***) indicate [statistical] significance at the 10%, 5%, and 1% level.

4.8.4 Summary statistics of the stochastic meta-frontier estimation

Table 4.6: Summary statistics of the stochastic meta-frontier estimation.

	Estimate	S.E.	T-Statistic
Meta-frontier			
$Constant$	0.0679	0.0033	20.6395
$\ln CH_4$	1.0040	0.0142	70.8158
$\ln N_2O$	0.0461	0.0149	3.0888
$\ln CO_2$	0.1834	0.0069	26.5320
$.5 \times \ln CH_4{}^2$	0.4395	0.1139	3.8572
$.5 \times \ln N_2O{}^2$	1.2780	0.1378	9.2753
$.5 \times \ln CO_2{}^2$	0.2428	0.0201	12.0759
$\ln CH_4 \times \ln N_2O$	-0.8298	0.1157	-7.1746
$\ln CH_4 \times \ln CO_2$	0.1944	0.0444	4.3745
$\ln N_2O \times \ln CO_2$	-0.3197	0.0485	-6.5867
t	0.0008	0.0007	1.1008
$.5 \times t^2$	0.0014	0.0005	3.1274
$\ln CH_4 \times t$	-0.0253	0.0048	-5.2176
$\ln N2_O \times t$	0.0288	0.0053	5.4602
$\ln CO_2 \times t$	-0.0060	0.0023	-2.6073
Environmental determinants			
$Constant$	-1.9292	0.3465	-5.5680
rA	2.0361	0.3424	5.9460
rB	2.0637	0.3419	6.0365
rC	1.8535	0.3407	5.4407
rE	2.1170	0.3421	6.1889
σ^2	0.1451	0.0059	24.6330
γ	0.9850	0.0017	589.5926

4.8.5 Percentage of farms per region belonging to the extensive or intensive class

Table 4.7: Percentage of farms per region belonging to the extensive or intensive class.

		Class 1 (extensive)	**Class 2 (intensive)**
Region a	(mountainous terrain)	52%	48%
Region b	(northern hill lands)	65%	35%
Region c	(tertiary terrain)	45%	55%
Region d	(fertile soil terrain)	50%	50%
Region e	(dry regions)	70%	30%

5 Study III – Using machine learning to identify heterogeneous impacts of agri-environment schemes in the EU: A case study

Disclaimer:

This is a pre-copyedited, author-produced version of an article accepted for publication in European Review of Agricultural Economics following peer review. The version of record (Stetter, C., Mennig, P. & Sauer, J.(2022). Using Machine Learning to Identify Heterogeneous Impacts of Agri-Environment Schemes in the EU: A Case Study. European Review Of Agricultural Economics) is available online at: `https://doi.org/10.1093/erae/jbab057`.

Acknowledgments

The authors thank two anonymous reviewers for their constructive comments on earlier drafts of this manuscript. The authors also thank Andrea Früh-Müller for the provision of the data on agri-environment schemes spendings at NUTS-3 level. This research was supported by the Bavarian State Ministry of Sciences, Research and the Arts in the context of the Bavarian Climate Research Network (bayklif).

5.1 Abstract

Legislators in the European Union have long been concerned with the environmental impact of farming activities and introduced so-called agri-environment schemes (AES) to mitigate adverse environmental effects and foster desirable ecosystem services in agriculture. This study combines economic theory with a novel machine learning method to identify the environmental effectiveness of AES at the farm level. We develop a set of more than 130 contextual predictors to assess the individual impact of participating in AES. Results from our empirical application for Southeast Germany suggest the existence of heterogeneous, but limited effects of agri-environment measures in several environmental dimensions such as climate change mitigation, clean water and soil health. By making use of Shapley values, we demonstrate the importance of considering the individual farming context in agricultural policy evaluation and provide important insights into the improved targeting of AES along several domains.

5.2 Introduction

The European Union's (EU) Common Agricultural Policy (CAP) has recently undergone its sixth major reform. While the EU's member states are about to adopt the European Commission's proposals regarding the post-2020 CAP (European Commission, 2018a,b,c), consensus prevailed among the main negotiators that environmental care, climate change action and the preservation of landscapes and biodiversity should be key elements of the new CAP. Especially the agriculture-induced loss of insects reported in recent studies (Ewald et al., 2015; Gossner et al., 2016; Ramos et al., 2018; Seibold et al., 2019) has spurred an intense public debate. But also indicators on soil erosion (Panagos et al., 2015), nitrate in groundwater, ammonia emissions (European Environment Agency, 2019b) and pesticide use (European Environment Agency, 2018) still do not, despite some positive trends, suggest an optimistic view. This situation is also a matter of concern given that today, at least 30% of the CAP's second pillar rural development spending must be allocated to investments in environmental and climatic sustainability, especially to AES. Voluntary AES in the context of CAP's second pillar have shown mixed success across Europe in terms of meeting environmental targets. Depending on the specific AES and the indicators under investigation, they have been found to be either beneficial (Batáry et al., 2015; Bright et al., 2015; Dadam & Siriwardena, 2019; Dal Ferro et al., 2016; MacDonald et al., 2012), ineffective (Bellebaum &

Koffijberg, 2018; Calvi et al., 2018; Granlund et al., 2005; Kaligarič et al., 2019; Kleijn et al., 2004), or even detrimental (Baer et al., 2009).

The question of how to adjust the design of AES to improve the delivery of a wide range of ecosystem services has been studied intensively (see e.g. Armsworth et al., 2012; Birge et al., 2017; Burton & Schwarz, 2013; Fuentes-Montemayor et al., 2011; Kuhfuss et al., 2016; Latacz-Lohmann & Breustedt, 2019; Latacz-Lohmann & Van der Hamsvoort, 1997; Westerink et al., 2017, 2014). More recently, a bundle of studies focused on (spatial) targeting of AES to improve the (cost-)effectiveness of such schemes (Desjeux et al., 2015; Früh-Müller et al., 2019; Langpap et al., 2008; Perkins et al., 2011; Uthes et al., 2010; van der Horst, 2007), which has often been neglected in past studies. It has been shown that both effectiveness and efficiency of AES increase if payments are well-tailored and well-targeted in space and time (Armsworth et al., 2012; Pe'er et al., 2020; Wätzold et al., 2016). This means, to increase the efficacy of their AE programs, policy-makers could specifically target farms where they expect a (large) positive treatment effect and adjust schemes where this is not the case. Typically, many of the above-mentioned analyses are biased toward an environmental and landscape perspective, and fail to provide a holistic picture of the targeting problem by ignoring farm-level effects. Studies that use farm-level data and classical statistical tools such as matching methods and/or difference-in-difference (DiD) estimators to assess the environmental effects of AES, on the other hand, only measure average treatment effects (ATEs) and fail to evaluate possible impacts at the individual level. Bertoni et al. (2020), for example, apply a conditional DiD coarsened exact matching procedure to estimate the average treatment effect on the treated (ATT) of three AES during 2007-2013. Similar DiD matching approaches were used by Pufahl & Weiss (2009), Chabé-Ferret & Subervie (2013), Arata & Sckokai (2016), Kuhfuss & Subervie (2018), and Uehleke et al. (2022).

In this paper, we demonstrate the usefulness of a novel machine learning (ML) approach to measure heterogeneous farm-level effects of AES participation. Besides the advantage of taking into account farm heterogeneity, ML methods such as the one used in this study can overcome multiple limitations of econometric and simulation models related to inflexible functional forms, unstructured data sources and explanatory variables (Storm et al., 2020). First studies that evaluate program participation based on ML methods have recently emerged in various fields ranging from personalized medicine to customized marketing. Within the field of agriculture and natural resources, Rana & Miller (2019) use the Causal Tree algorithm developed by Athey & Imbens (2016) to assess the impact of two community forest management policies on vegetation in the Indian Himalaya. Deines et al. (2019) use the generalized random forest (GRF) algorithm to study the effect of conservation tillage practices in the US Corn Belt based on satellite-derived data. Further applications include Carter et al. (2019) using GRF to evaluate rural development programs in Nicaragua, and Mullally & Chakravarty (2018) applying LASSO to study the effects of rural business development programs on production and productivity in Nicaragua. Only recently, Miller (2020) used causal forests to analyze the impact of quotas on fisheries' catches around the world.

Following these novel research approaches, we seek to overcome several limitations of previously used econometric impact evaluation methods by making use of an innovative ML algorithm to assess the heterogeneous effects of agri-environmental measures. We demonstrate the merits of this approach for the case of the German Federal State of Bavaria in the 2014-2020 CAP programming period. In line with the environmental priorities for the 2014-2020 CAP Rural Development pillar defined by the European Commission (2013), which mainly target biodiversity enhancement, improvement of water and soil quality and greenhouse gas emission reduction, we develop comprehensive indicators for each sub-goal and test the heterogeneous AES efficacy for these indicators. While the success of AES largely depends on a large variety of individual farm characteristics as well as on the biophysical and institutional context (Dupraz & Guyomard, 2019), legislators cannot take account of all individual characteristics of the eligible farms when designing and targeting AES, e.g. to avoid discrimination, which inevitably leads to inefficiencies (Dessart et al., 2019; Dupraz & Guyomard, 2019). Given the capability of our research approach to obtain farm-specific treatment effects, we evaluate several dimensions according to which policy-makers might target specific farm groups to improve the efficacy of their AES (location, size, farm typology and yield potential) by means of Shapley values, a model-agnostic concept stemming from the interpretable ML literature.

First, farm location is given special attention in this regard. For instance, Pelosi et al. (2010), Matzdorf et al. (2008) and Früh-Müller et al. (2019) find spatial inefficiencies in multiple environmental dimensions such as soil health, water quality and habitat fragmentation. They argue that spatial targeting of AES could strongly improve their environmental efficacy. Furthermore, the spatial dimension of AES is emphasized by Desjeux et al. (2015), Dessart et al. (2019) and Coderoni & Esposti (2018). Second, farm typology is considered as an important driver of AES effectiveness (Coderoni & Esposti, 2018; Westbury et al., 2011). Given their farming context, most farms are bound to specific technologies, which is why increasing the uptake of farm groups belonging to certain farm typologies is likely to increase AES efficacy. For instance, Herrero et al. (2016) find that greenhouse gas mitigation potentials are particularly large in the livestock sector. Thus, targeting specific farm types might lead to higher AES efficacy. Third, farm size, e.g. expressed by farmed area, as part of farm characteristics should also affect the AES treatment effect. In terms of extensification, Wuepper et al. (2020) suggest that small farms cannot easily afford to take land out of cultivation compared with larger farms. Hence, we would expect a positive impact of farm size on AES effectiveness. Other studies, such as Coderoni & Esposti (2018) and Westbury et al. (2011) do not find that farm size affects AES effectiveness regarding GHG emissions and extensification, respectively. Since farm size is already used as a target dimension within the Single Payment Scheme of the CAP (Salhofer & Feichtinger, 2020), this might also be a good option for AES. Fourth, another important contextual factor is yield potential. Legislators have started to realize that targeting farms according to their yield potential might increase their AES-effectiveness (ART, 2019). To assess how policy-makers can make use of the above-mentioned contextual variables to improve AES-effectiveness, we answer the following two questions:

1. How do location, size, farm typology and yield potential affect the impact size of AES?
2. How can legislators use location, size, farm typology and yield potential to target specific farm groups?

This allows us to draw conclusions as to how legislators can improve the effectiveness of AES by better targeting their policy measures considering farm-level characteristics.

The remainder of this article is structured as follows. In Section 5.3, we provide some background on AES and describe the conceptual underpinnings of the study. Section 5.4 provides information on the data used in this study, while Section 5.5 refers to the analytical framework. In Section 5.6, we describe and discuss the empirical findings and their policy implications. The final section 5.7 summarizes and concludes the study, also providing promising directions for further research.

5.3 Conceptual framework and background

5.3.1 AES description

Our case study region, the Federal State of Bavaria offers a range of AES as part of its 2014-2020 rural development program (RDP), which was extended until 2022 due to ongoing CAP negotiations. The individual measures are grouped into two RDP subprograms, the Nature Conservation Program (Vertragsnaturschutzprogramm, VNP), and the Bavarian Cultural Landscape Program (Bayerisches Kulturlandschaftsprogramm, KULAP). While VNP schemes are only to be implemented in pre-defined areas of high nature value, KULAP measures are generally not directly linked to specific areas and applicable in entire Bavaria.

There are up to 42 individual KULAP schemes and 36 VNP schemes that are offered within the current RDP. All of the schemes offered in 2014, the year we focus on, are action-based, i.e. the scheme payments are linked to certain farming requirements. The superordinate goals of both VNP and KULAP refer to maintaining/improving biodiversity, water and soil quality and mitigating climate change. Each category is related with a number of AES, with multiple schemes being assigned to several environmental goals. Such a multi-target approach is related to interdependencies between non-marketed goods and services. An AES restricting the use of mineral fertilizer, for example, contributes simultaneously to water protection and greenhouse gas emission reduction in the absence of leakage effects. Sometimes, however, the impacts of certain AES positively affect one target and adversely affect another target (Knudson, 2009). This relation is linked to the Tinbergen Rule (Tinbergen, 1956), which states that efficient policy requires at least as many policy instruments as there are targets, i.e. each instrument should address a single goal (Huber et al., 2017).

The existing mix of (agri-environment) measures and environmental goals complicates impact evaluations. First, farmers often participate in several AES at the same time. If, as in our case, only a variable indicating if a farmer participates in any scheme is given, but no information on the exact type(s) of scheme(s), possible effects cannot be traced back to a certain sub-scheme. And even if this information was available, it would be difficult to unambiguously link effects to specific schemes given their multitude of goals and combination possibilities (Chabé-Ferret

& Subervie, 2013). We address this issue in our approach by focusing on the overarching aims of the Bavarian agri-environment programs of improving biodiversity, soil and water quality and reducing greenhouse gas emissions. These goals apply in the entire federal state. Our analysis allows to identify regions and farm types that respond strongly and weakly to AES participation in terms of environmental outcomes. It also provides the basis for linking effect sizes to individual scheme uptake. In this regard, our approach is in line with existing studies on AES impact evaluation such as Pufahl & Weiss (2009), Arata & Sckokai (2016), and Mennig & Sauer (2020).

5.3.2 Production possibilities and farming context

To understand the impact of AES on the environmental performance of farms, it is useful to think about how they affect farms' production possibilities. A standard approach is to assume that all firms share the same production possibilities (Chambers, 1988). The production possibilities depend on the available resource or input bundle. Introducing a binding action-based AES typically means limiting the resource bundle and thus also limiting the production possibilities. Given the multi-functional nature of farming, this affects both agricultural (e.g. crop and livestock) outputs as well as ecosystem services (e.g. soil formation, biodiversity, or climate change) through their joint production (Wossink & Swinton, 2007).

In agriculture, the assumption that production possibilities are the same for all farms is quite unrealistic for a number of reasons (Tsionas, 2002). For instance, the available resource bundle and input intensity are at least partially exogenously determined by the production or biophysical environment (weather, topography, soil quality etc.), which is defined as features that are physically involved in the production process (O'Donnell, 2016). Furthermore, given the stationary nature of farming regarding its location, the institutional environment as well as factor (e.g. capital, labor, and land) and output market imperfections determine farms' point of production. This results in farm-specific factor endowments, cultivation plans and yields. For these reasons, the production possibilities of farms are usually bound to specific technologies, which cannot be easily switched (e.g. crop farming vs. livestock farming, or grassland vs. arable farming). Finally, the point of production depends also on farmer-related characteristics. This includes both socio-demographics as well as behavioral factors (Dessart et al., 2019).

Bearing the heterogeneous nature of production (possibilities) in mind, Figure 5.1

provides four stylized cases describing potential scenarios farmers face when deciding to participate in an AES. To avoid undue complexity, only two outputs are considered, namely composite agricultural goods and environmental services. Figure 5.1 depicts several production possibility frontiers (PPFs), which illustrate the combinations of outputs that the farm can produce.

All points on or beneath the curve are feasible. The optimal point of production is where the iso-revenue line (IR), which depends on the marketed output and its price, is tangent to the PPF. Since there is no price explicitly assigned to environmental services, the IR is horizontal. Here, a complementary-competitive relationship between agricultural and environmental outputs is assumed such that (at least) the range close to the Y-axis is convex (Sauer & Wossink, 2013a; Wossink & Swinton, 2007).[1]

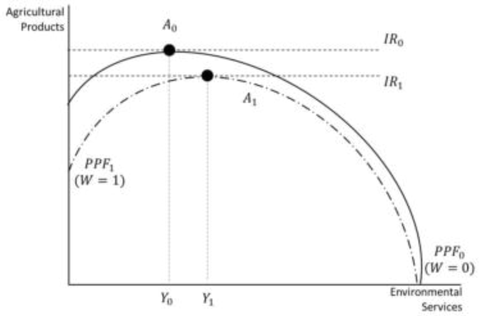
Action-based AES are part of the production environment and usually require certain behaviors that restrict the available PPF of a farm (see Section 5.3.1). Hence, a farm faces a decision between two potential PPFs, whose shape and location are determined by the above-mentioned farm-specific contextual factors. PPF_0 is the PPF with no AES restrictions. PPF_1 is the PPF with AES restrictions. In the case of Fig. 5.1a, the farm decides to produce either at point A_0 (no AES) or A_1 (AES). Hence, the farm foregoes agricultural output[2] for environmental output (Y) when participating in the AES. The difference between the two potential environmental outputs Y_1 and Y_0 is the treatment effect of participation.

Fig. 5.1b shows a situation of an inefficient farm producing beneath the potential PPFs. Participating in the program does not change its point of production and therefore $Y_1 = Y_0$. The direction to the PPF is also context-specific and endogenously determined by the farm (Färe et al., 2013). Fig. 5.1c depicts a situation, in which the AES does not shift the PPF and participation in the program does not change the point of production A. In Fig. 5.1d the AES changes the PPF such that the $Y_1 - Y_2$ is negative, which means an adverse participation effect.[3] Scenarios 2 and 3 describe a situation, in which farms profit from a windfall

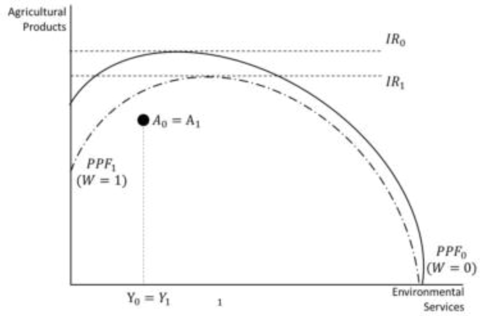
[1] A typical example where a marginal increase in ecosystem services leads to increased agricultural output would be the cultivation of cover crops, which helps avoid soil erosion while at the same time enhancing soil fertility. The same reasoning as presented here can also easily be applied to supplementary and competitive relationships (see e.g. Sauer & Wossink, 2013a). Appendix 5.8.1 illustrates two straightforward versions of these cases.

[2] We assume the program compensates adequately for this loss. Otherwise, a rational farmer would not sign up for the program.

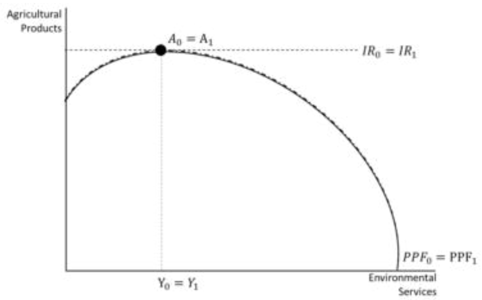
[3] This case is most likely if there is a negative trade-off effect of an AES in terms of different environmental outcomes. E.g. a measure has an additional effect on land use diversity but



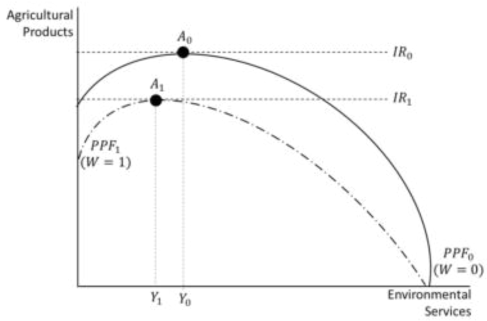
(a) Desired response to AES participation. Efficient farm increases production of desired environmental output from Y_0 to Y_1 as production possibility frontier shifts under AES participation.



(b) No output response to AES participation. Inefficient farm is not directly affected by a shift of its production possibility frontier under AES participation.



(c) No output response to AES participation. AES does not shift the production possibility frontier of the efficient farm.



(d) Adverse response to AES participation. Efficient farm decreases production of desired environmental output from Y_0 to Y_1 as production possibility frontier shifts under AES participation.

Figure 5.1: Stylized cases reflecting the potential impact of AES participation under heterogeneous production possibilities with one agricultural and one positive environmental output. Action-based AES change the resource and input bundle of farms, thus changing farms' production possibilities. Hence, farmers face two potential production possibilities, of which only one can be realized. Depending on individual farm, institutional, and environmental characteristics, the shape and location of the PPF varies across farms. As no price is assigned to environmental outputs, iso-revenue lines are horizontal. Under the assumption of a fixed resource and input bundle, an efficient farm produces at the point where the iso-revenue line (IR) is tangent to the PPF.

effect, i.e. they receive an environmental subsidy without having to adjust their agricultural practices. Scenario 4 can be seen as a worst-case-scenario as farms receive compensation although their environmental service declines. Scenario 1 represents the expected effect by policy-makers. As the AES is not designed to match the individual production environment, all four cases can occur depending on the heterogeneous farming context (see also Sec. 5.4.3).

The same line of argument concerning the production possibilities and farming context carries over to the farm's decision to enter the program. If the opportunity cost of providing ecosystem services is covered by the program's compensation, we expect a farm to enter given their farming context (Sauer & Wossink, 2013a). This context determines the provision of environmental services through altering opportunity costs, i.e. the revenue foregone by providing non-marketed goods and services. Consequently, for some farms the payments for specific AES, which are generally the same for all farms, will be too low to participate, while others might not face opportunity costs as even in the absence of the scheme their farm management would have been the same. Generally, the farming context determines if the opportunity cost of program participation is covered by the AES compensation and hence if a farm enters the program.

5.3.3 Conditional average treatment effects

Section 5.3.2 points out that the treatment effect of AES is expected to vary across farm households. Although acknowledged by many previous studies on the subject, most of them could only estimate average effects on the basis of traditional statistical methods. Our approach, however, is based on the conditional average treatment effect (CATE) that allows to obtain individualized AES treatment effects.

Having two potential outcomes Y^0 and Y^1 (see Figure 5.1), we embed the problem into the Rubin causal model (Neyman, 1923; Rubin, 1974). Suppose a set of i.i.d. farm households $i = 1, \ldots, n$, for which we observe (X_i, Y_i, D_i), where $X_i = x \in \mathbb{R}^p$ is a vector of p features[4], describing the individual farming context and containing all determinants of Y^0 and Y^1 as well as the determinants of the participation decision.[5] $Y_i \in \mathbb{R}$ is the outcome variable of interest (e.g. an indica-

adversely affects greenhouse gas emissions.

[4]The term feature corresponds to "coavariate" in the traditional econometric terminology.

[5]By this formulation, we allow for the fact that all variables in the model might possibly be confounding factors. Thus, we avoid making *a priori* assumptions as to which variables are

tor reflecting environmental performance), and $D_i \in \{0,1\}$ is the policy dummy for participation and non-participation in AES. Given the potential outcomes Y_i^0 and Y_i^1, for each farm i that is (uniquely) characterized by its feature vector x, we wish to estimate the CATE: $\tau(x) = \mathbb{E}\left[Y_i^1 - Y_i^0 \mid X_i = x\right]$. However, following Holland (1986), it is impossible to observe the effect for more than one treatment on a subject. Hence, we can only observe realization $Y_i = Y_i(D_i)$. Without further assumptions, it is impossible to identify the CATE $\tau(x)$. Therefore, we invoke the conditional independence assumption (Rubin, 1977), i.e. D_i is independent of unobservable features conditional on X_i: $Y_i^1, Y_i^0 \perp\!\!\!\perp D_i \mid X_i$. Furthermore, we assume common support to rule out perfect predictability of program participation, i.e. individuals with the same X have a positive probability of being both participants and non-participants: $0 < P(D_i = w \mid X) < 1$. We then define the propensity score $e(x) = \mathbb{P}\left[D_i = 1 \mid X_i = x\right]$ for the probability of being assigned to the treatment conditional on X, and $m(x) = \mathbb{E}[Y_i = y \mid X_i = x]$ for the expected outcome conditional on X. Given the aforementioned assumptions, and based on the findings by Robinson (1988) and Chernozhukov et al. (2018), Athey et al. (2019) argue that the CATE can be identified by the simple outcome model $Y_i = \tau(x)D_i + m(x) + \epsilon_i$. Transforming this into a residuals-on-residuals regression (Chernozhukov et al., 2018), we obtain the following estimator:

$$Y_i - \hat{m(x)} = \tau(x)\left(D_i - \hat{e(x)}\right) + \epsilon_i \tag{5.1}$$

where ϵ_i is a random error term. One advantage of using this residual-on-residual approach is that it makes the parameter estimate (τ) insensitive to small errors in the formulation of $m(x)$ and $e(x)$, thus improving its robustness (Athey et al., 2017; Chernozhukov et al., 2018). Furthermore, it is a "doubly robust" estimator. Doubly robust estimators are unbiased if one specifies at least one of the nuisance models correctly (i.e. the treatment $e(x)$ and outcome model $m(x)$) (Chernozhukov et al., 2018). Hence, this estimator is effectively a debiasing routine, which should yield a robust parameter estimate of the CATE τ under the given assumptions.

5.4 Data and variable description

In our analysis, we mainly rely on farm accountancy data for the German federal state of Bavaria. Located in the southeast of Germany, Bavaria belongs to

confounding factors.

the core regions of agricultural production within the EU. Its heterogeneous natural conditions are well-suited for various agricultural production systems such as crop farming, intensive and extensive dairy farming, pig and cattle fattening and breeding, poultry farming, vegetable farming, orcharding, hop production and viticulture. This heterogeneity of farming systems represents to some extent the European agricultural sector and is reflected by a broad variety of Bavarian AES. We chose to analyze data from 2014 as the first year of the then new CAP period. Our data are part of the European Farm Accountancy Data Network (FADN) with a sample size of 2758 observations. We do not restrict the dataset to specific farm types. However, organic farms are excluded from the analysis due to their distinctly different farming approach compared to conventional farms. The sample is stratified with respect to farm location, size classes, and specialization of the farms. In addition to financial records, the dataset contains information about, for example, the cultivation plan, yields and socio-economic information such as the educational level of the farm manager, the number of household members or the on-farm labor structure. We match the farm accountancy data to official agricultural support data containing information about farm-specific scheme participation as well as to secondary data at the county-level to retrieve further information on the socioeconomic, spatial and structural environment of the farms.

5.4.1 AES indicator

For our empirical analysis, we use a binary treatment variable, which takes on a value of 1 if a farm participated in an agri-environmental scheme in 2014.[6] Farms that did not participate were assigned a value of 0 for the treatment variable D. For Bavaria, we find that 1641 farms participated in an AES in 2014, while 1117 did not. As outlined in Sec. 5.3.1, we choose a generic binary AES indicator for two reasons. First, our data does not contain detailed information on individual sub-schemes. Second, even with this information, it might be impossible to unambiguously determine CATEs for individual sub-schemes because they are inherently inseparable (Heiler & Knaus, 2021).

5.4.2 Environmental indicators

In order to assess the environmental performance of the sample farms, we make use of four comprehensive, well-established environmental farm-level indicators to

[6]Detailed information about AE schemes in Bavaria can be found in Appendix 5.8.2.

properly evaluate the four domains of more environment-friendly farming practices, namely soil and water health, biodiversity, and GHG mitigation.

First, within the soil/water domain and following studies such as Uehleke et al. (2022) and Arata & Sckokai (2016), we select fertilizer and pesticide intensity as environmental outcome variable, which we define as expenses in € per hectare of land. Second, we seek to assess farm-level (bio-)diversity by means of the Gini-Simpson diversity index (g_i) containing all managed land use types in the dataset and calculated by the following formula:

$$g_i = \left(1 - \sum_k s_{ik}^2\right) \times 100 \tag{5.2}$$

where s_{ik} stands for the share of land-use type k on farm i. The higher the value for the Gini-Simpson index is, the greater is the land use diversity of a farm. Studies show that the more heterogeneous landscapes are, the higher is their provision of a multitude of environmental services such as enhanced soil nutrient cycling, mineral retention, regulation of pests and pathogens, as well as an improvement in pollination and water quality (see e.g. Brussaard et al., 2007; Smukler et al., 2010; Tomich et al., 2011). There is also some evidence that they are a key determinant of biodiversity (Benton et al., 2003). Third, regarding the climatic impact of agriculture, we use the farm-level carbon footprint index developed by Baldoni et al. (2017). Their greenhouse gas inventory approach exploits relevant activity data regarding various emission sources, which is contained in the farm accountancy dataset. These activity data are then multiplied with the respective regional emission factors contained in Haenel et al. (2018). This method closely follows the recommendations of the Intergovernmental Panel on Climate Change (IPCC) and allows for a farm-level assessment of the three most important greenhouse gases in agriculture, namely methane, nitrous oxide and carbon dioxide stemming from various farming activities (compare Baldoni et al., 2017; Coderoni & Esposti, 2014, 2018).[7] Descriptive statistics for the four indicators can be found in Table 5.1. While we find lower levels of fertilizer and pesticide intensities as well as a higher diversity index on average for the participating farms, they unexpectedly emit more GHG emissions than the control group. One explanation for this could be that treated farms are on average larger (in terms of whole-farm value

[7] To combine GHG emissions in one indicator, methane and nitrous oxide emissions were converted to CO_{2eq}. To that end, N_2O and CH_4 quantities were multiplied by their respective global warming potentials (34 and 298, respectively) as per the IPCC's Fifth Assessment Report (IPCC, 2013), considering the inclusion of climate carbon feedback and a 100-year time horizon.

added and farm land). Other than the first three indicators, GHG emissions are measured in absolute numbers[8], which is why this pattern might occur.

Table 5.1: Descriptive statistics - ecological responses.

		Treated (N=1677)		Untreated (N=1081)		Entire sample (N=2758)	
Domain	Indicator	Mean	SD	Mean	SD	Mean	SD
Soil/water	Fertilizer intensity (Euro/ha)	186.69	90.73	205.03	95.16	194.12	92.97
Soil/water	Pesticide intensity (Euro/ha)	120.19	99.62	121.13	105.46	120.57	102.01
Biodiversity	Gini-Simpson index (0-100)	67.23	21.27	63.65	19.14	65.78	20.51
Climate	GHG emissions (t CO_{2eq})	469.39	370.8	411.01	334.21	445.75	357.52

5.4.3 Features

As outlined in Section 5.3.2, the effect of the participation in AE schemes depends on a multitude of factors. We identified the following domains, according to which the treatment effect may vary for their influence on farms' production possibilities:

- Resource bundle and input intensities (e.g. Tsionas, 2002)
- Output bundle (e.g. Sauer & Wossink, 2013a; Wossink & Swinton, 2007)
- Farm and farmer characteristics (e.g. Dessart et al., 2019)
- Biophysical environment (e.g. Desjeux et al., 2015; O'Donnell, 2016)
- Institutional and market environment (e.g. Landini et al., 2020)

The individual heterogeneity domains are described by a rich set of observable covariates, which are depicted in Table 5.2.[9] Due to the strong nonlinear mapping and adaptive prediction functionality of random forests (RFs), we do not have to arbitrarily aggregate covariates. This is a clear advantage of the ML approach compared to more traditional parametric models. The richness of the variables in our model allows us to capture the real-world complexity of farms very well, which is likely to influence both the propensity of participating in an agri-environmental

[8]This is because the absolute atmospheric pressure must be reduced to be effective. In contrast, pesticides and fertilizers, have mostly a more local effect, which is why they are measured per unit of land (i.e. ha).

[9]Descriptive statistics of the whole feature set can be obtained from the authors upon request.

scheme as well as the effect size itself. Compared to more traditional econometric techniques, this is a clear strength of the machine learning algorithm.

Input intensities and the farm-specific resource bundle are described by a combination of land use, labor, materials and capital. Furthermore, our empirical strategy allows to include the complete cultivation plan and livestock count of each farm. Next, the output bundle is described by a total of ten different output variables. Farm and farmer characteristics include, among other variables, farm type, decoupled subsidies, value added, farmers' age and education as well as yield data approximating farmers' productivity levels and management capacities. The primary proxy for the locational setting of the farm is described by a county indicator variable. Furthermore, the biophysical environment is further described by a yield index unit describing the farm-level soil quality and yield potential for each farm and information on the altitude. The institutional and market environment is further approximated, e.g., by county-level land rental prices (land market), unemployment rate and population density (labor market). As stated earlier, special attention will be given to the four targeting dimensions, namely farm size, i.e. total land, farm type[10], yield index unit as well as farms' location (approximated by county affiliation).

The fact that the analysis is bound to cross-sectional data gives rise to two potential sources of endogeneity. First, we cannot control away time-constant unobserved heterogeneity through fixed or random effects. We address this issue in Section 5.5.3. Second, looking at Table 5.2, many covariates describing the individual production possibilities might already be influenced by the treatment itself, thus inflicting post-treatment bias by controlling away for the consequences of treatment (King & Zeng, 2006; Montgomery et al., 2018; Wooldridge, 2005). To shut this feedback path between treatment and controls, we use long-term average values from the previous AES period (2007-2013) to describe the farming context for all covariates reflecting the resource and input bundle, the output bundle, and farm characteristics, which might all be directly affected by AES participation itself.

The implementation of the causal forest is designed for complete data. As there

[10]Specialized farms (dairy, pig and crop) are assigned to the respective farm type if the output share of their characteristic produces exceed 66% in total revenues (milk, cattle, poultry, fattening pigs, grains). As for mixed farms (i.e. crop-livestock systems), no primary product accounts for more than 66 % of total revenues.

Table 5.2: Description of the predictor space for the estimation of the causal forest.

Heterogeneity Domain	**Predictors**
Resource bundle & input intensity	
– Land use	Total land (ha), rented land (ha), own land (ha), arable land (ha), grassland (ha), share rented land (0-1), share grassland (0-1)
– Labor (man-work units)	Total on-farm labor, family labor, hired labor, labor intensity (€/ha)
– Materials and capital (€)	Seed expenditure, feed expenditure, capital expenditure, capital intensity (€/ha), feeding intensity (€/ha)
– Cultivation plan (ha)	Winter wheat/spelt, spring wheat, durum wheat, rye, winter barley, spring barley, oat, winter cereal mixture, spring cereal mixture, grain maize, corn cob mix, triticale, other cereals, field beans, feed peas, other feed legumes, other legumes, winter canola, spring canola, sunflowers, soybeans, linseed, other oilseeds, energy corn, energy cereals, energy legumes, energy oilseeds, energy beets, potatoes, sugar beet, cabbage^{+}, leafy vegetables^{+}, fruit vegetables^{+}, asparagus^{+}, other tubers^{+}, legume vegetables^{+}, other vegetables^{+}, tobacco, grass seeds, other seeds^{+}, minor plants (e.g. medicinal plants), other energy plants, other renewable resources, ground ear maize, feed root crops, clover, cover crops, temporary grassland, permanent grassland, alpine pasture, cereal forages, hops, set-aside land, set-aside land (minimum 10 years), fallow
– Livestock count	Light horses, heavy horses, male beef, dairy cows, suckler cows, calves, heifers, male cattle, weaners, fattening pigs, sows, boars, sheep, pullets, laying hens, broilers, poultry
Agricultural output bundle (€)	Cereals, canola, potatoes, sugarbeet, other plants, milk, pigs, cattle, livestock total, crop total
Farm characteristics	Farm type, whole farm value added (€), value added per ha (€/ha), Full time farm (yes/no), age (years), agricultural education (none, low, high), Milk yield (liters/cow), potato yield, winter wheat yield, spring wheat yield, grain maize yield, canola yield, general pulses yield, bean yield, fodder plant yield, rye yield, winter barley yield, spring wheat yield, oat yield, triticale yield, pea yield, sugarbeet yield, silage maize yield
Biophysical environment	Administrative units (*counties*), yield index unit, altitude (<300m, 300-600m, >600m)
Institutional environment and markets	Administrative units (*counties*), GDP per capita (€), gross value added in agriculture (mio. €), unemployment rate (%), population density (habit./km^2), land rental price (€/ha)

$^{+}$ Field cultivation

are very few missing values in the dataset, we impute the missing data points by means of fully conditional specification using Breiman's RFs as described in Doove et al. (2014).

5.5 Analytical framework

5.5.1 Using causal forests to estimate the CATE

Following the residual-on-residual approach from Sec. 5.3.3, to obtain the CATE estimate $\tau(x)$ (Eq.5.1), both environmental outcome $m(x)$ and participation probability $e(x)$ must be predicted in a first step. One possibility to obtain such estimates would be to estimate a parsimonious parametric model. However, this model would likely be inappropriate in high-dimensional settings[11]. For that reason, Athey et al. (2019) suggest RFs to estimate $m(x)$ and $e(x)$ and finally also $\tau(x)$.

RFs, a concept developed by Breiman (2001), are basically an ensemble of regression or classification trees (CART), which are grown based on recursive partitioning such that the feature space is divided into binary nodes according to an optimality criterion (e.g. many standard regression tree implementations split by minimizing the in-sample prediction error of the node (Breiman et al., 2017)) until the final nodes (aka leaves) contain a number of observations greater than a given minimum. The average outcome of such a leaf is then the prediction for the observations contained in that leaf. RFs make predictions in the form of an average across predictions $b = 1, \ldots B$ of such CARTs, each of which is grown on a training sample, i.e. a random subsample of the data. Based on that, Athey & Imbens (2016) and Wager & Athey (2018) formally establish asymptotic normality for regression trees and RFs through *honest splitting* of trees, i.e. the training sample is split into two parts, one part is used to train the tree and the other part is used to predict the outcome of interest.

Athey & Imbens (2016) demonstrated how treatment effects could be computed based on regression trees by means of an adjusted splitting rule by the finding that squared-error minimizing splitting is equivalent to maximizing the heterogeneity across child nodes. Wager & Athey (2018) build upon these findings and introduce *causal forests* that average the tree-based effects for each individual over a large

[11] If we assume that the true relationship between, e.g. outcome and features is rather complex and contains many features, linear models usually fail to grasp high-dimensional interactions and nonlinearities and are prone to model misspecification and variance inflation.

set of trees. Athey et al. (2019) generalize these findings to a broader context of estimation methods, in that they regard RFs not as an ensemble method (aka averaging the results of multiple trees) but as an adaptive kernel method, e.g. some outcome Y_i could be predicted by means of $\widehat{f}(x) = \sum_{i=1}^{n} \alpha_i(x) Y_i$, where $\alpha_i(x)$ is a data-adaptive-kernel measuring how often the i-th observation falls in the same leaf as a test point x. The causal effect specific similarity weights $\alpha_i(x)$ can be obtained by means of a *causal forest* based on trees that greedily optimize for treatment effect heterogeneity across child nodes based on a local moment condition.[12] A more detailed description of the estimating strategy can be found in the Appendix 5.8.3. The weights are formally defined as:

$$a_i(x) = \frac{1}{B} \sum_{b=1}^{B} \frac{\mathbf{1}\left(\{X_i \in L_b(x), i \in S_b\}\right)}{\mid L_b(x), i \in S_b \mid} \tag{5.3}$$

where $L_b(x)$ is the leaf of the b-th tree that contains the test point x and S_b denotes the subsample used to grow the b-th tree. Athey et al. (2019) show that after growing a *causal forest* to obtain the forest weights $\alpha_i(x)$, the locally weighted estimator for the treatment effect $\hat{\tau}(x)$ is

$$\hat{\tau}(x) = \frac{\sum_{i=1}^{n} \alpha(x) \left(\widetilde{D}_i - \bar{D}_\alpha\right) \left(\widetilde{Y}_i - \bar{Y}_\alpha\right)}{\sum_{i=1}^{n} \alpha(x) \left(\widetilde{D}_i - \bar{D}_\alpha\right)^2} \tag{5.4}$$

where $\widetilde{Y}_i = Y_i - \hat{m}(x)^{oob}$ and $\widetilde{D}_i = D_i - \hat{e}(x)^{oob}$[13], $\bar{D}_\alpha = \sum_{i=1}^{n} \alpha(x) \widetilde{D}_i$ and $\bar{Y}_\alpha = \sum_{i=1}^{n} \alpha(x) \widetilde{Y}_i$. From (5.4) it becomes apparent that the heterogeneity of the conditional treatment effect fundamentally stems from the causal forest weights $\alpha_i(x)$.[14]

What is more, by using an orthogonalized causal forest (see Appendix 5.8.3 Eq. 5.8) in the spirit of Eq.5.1 and obtaining estimates for the propensity scores $\hat{e}(x)^{oob}$, the estimator (5.4) is robust to potential confounding effects. This makes the presented procedure well-suited to analyze observational data.

Athey et al. (2019) show that valid confidence intervals for causal forest estimates can be obtained by means of the 'bootstrap of little bags method', where basically

[12] Athey et al. (2019) found that the *causal forest* (Wager & Athey, 2018) can be seen as a special case of the GRF for a binary treatment variable. Therefore, the terms causal forest and GRF are used interchangeably throughout this thesis.

[13] *oob* denotes out-of-bag predictions, i.e. these predictions are generated by using only the portion of trees that do not have that data point in the respective subsample used to generate the predictions.

[14] Basically, these weights could also be computed using traditional k-NN estimates. However, k-NN is limited in the sense that it does not distinguish with respect to feature importance. As RFs are data-adaptive and thus prioritize high-signal features, it is better-suited to yield precise weights in a high-dimensional feature space (Athey et al., 2019).

small groups of trees are trained and their predictions are then compared within and across groups to estimate the variance. For a more technical description of the method, see Sexton & Laake (2009).

5.5.2 Model specification

In a first step, we fit a propensity forest to estimate the predicted propensity scores $\hat{e}(X_i)^{oob}$ of each farm i. We specify the number of trees to 5000 in order to obtain stable estimates in the sense that they yield the same predictions if we grow forests of the same size on the same data set. We perform parameter tuning on this forest to improve overall model performance (James et al., 2021), i.e. the minimum number of observations in each tree leaf, the fraction of the data used for the subsample to build each tree, the number of variables tried for each split, as well as split balance parameters are chosen by means of cross-validation. As mentioned in Section 5.5.3, by using a high-dimensional set of predictors, we are confident to obtain reliable propensity scores that largely capture background differences between participants and non-participants and serve as proxies for features that were not included (Rana & Miller, 2019) such that the unconfoundedness condition appears to be satisfactorily plausible in this setting.

Second, we estimate a separate regression forest for every environmental indicator to obtain $\hat{m}(X_i)^{oob}$. Again, we determine the hyper-parameters of the forest through tuning and train 5000 trees. Third, given $\hat{e}(X_i)^{oob}$ and $\hat{m}(X_i)^{oob}$, we can train a causal forest to obtain heterogeneous treatment effects (HTE, $\hat{\tau}$) for each environmental outcome. As this forest yields the final estimates of interest, we are more stringent in terms of the prediction stability and fit 100,000 trees for each environmental indicator. By doing this, we guarantee that the excess error – measuring the stability of our estimates – is negligibly small (Wager et al., 2013). Furthermore, as before, we use hyperparameter tuning using cross-validation to improve the performance of the algorithm.

5.5.3 Latent confounders and omitted variable bias

One major criticism of the identification strategy presented in Section 5.3.3 is undoubtedly the selection-on-observables assumption, i.e. the heterogeneous treatment effect is only identified if all relevant confounders are observed by the researcher (see also a graphical visualization in Figure 5.8 in the appendix). Otherwise, the estimates will be biased due to unobserved omitted variables that are correlated with both treatment and outcome (DiPrete & Gangl, 2004). Here, we rest upon recent advances in the causal ML literature (Bennett & Kallus, 2019; Kallus et al., 2018; Louizos et al., 2017; Wang & Blei, 2019) and make the case that, by using RFs, we may tackle endogeneity bias stemming from unobserved heterogeneity although we do not include all potential confounding factors directly.

The reasoning behind this is as follows (see also Figure 5.9 in the appendix). The nonlinear, highly-complex combination of a high-dimensional set of the observed potential confounding features X serves as an approximation of the unobserved confounding factors and is able to represent the latent covariate space to a certain degree, which remains unobserved to the researcher. One classical example for a latent confounder in the context of AES is farm managers' attitude toward the environment, affecting both the participation decision as well as environmental outcome.[15] Through the nonlinear, high-dimensional combination of a large number of observed proxy features (X) such as farming conditions (e.g. agri-climatic regions, yield potential, altitude), county-level settings, farm type, farm size, land and capital use, labor structure, education and productivity indicators such as milk yield (compare Section 5.4)[16], we argue that the causal forest through its complex structure is able to capture (a lot of) the variation coming from this unobserved confounder space.[17] RFs are very effective at uncovering such latent structures (similar to neural networks). Such a representation is not possible with conventional regression techniques, which are only able to assess an often linear, low-dimensional feature space, and which therefore are not able to approximate the latent space sufficiently.

The assumption that causal forests are able to approximate well omitted variables might thus be one response to tackle the unconfoundedness condition. Note, in order to effectively mitigate omitted variable bias, we rely on the assumption that all relevant information is latently contained in our observed data. If there was a completely different group of confounding variables that are not contained in the included confounders, our estimates might still be biased (see also Figure 5.10 in

[15] Another example for such a confounder would be managerial ability.

[16] A multitude of studies found a close association between environmental attitude and observed characteristics (Borges et al., 2015; Farr et al., 2018; Featherstone & Goodwin, 1993; Prokopy et al., 2019, e.g.). In line with this, Austin et al. (2001) find that (environmental) attitudes and managerial ability are manifested in (observable) management practices.

[17] In practice, this means that if two observations end up in the same leaf of a RF that splits several times on the above-mentioned features,these two observations have a (nearly) identical attitude toward the environment. We assume that miscellaneous variation in the latent variable is idiosyncratic and has low to no signal.

the appendix). To test the sensitivity of this latent variable assumption, we suggest a range of robustness checks testing the stability of our model to omitted variable bias coming from unobserved confounding factors. These imply several placebo, leave-p-confounders-out tests and the simulation of additional confounders under different correlation structures. A detailed description of the sensitivity checks can be found in the appendix (5.8.6).

5.6 Empirical results and discussion

5.6.1 AES program uptake and indicator prediction

The trained propensity forest yields plausible propensity score estimates (Figure 5.2, panel A). The scores are bounded between 0.27 and 0.86. We do not find any propensity that is very close to 0 or 1. This is still true if we regard the uncertainty of our estimates by including their 95% confidence intervals (Figure 5.2, panel b)). To be consistent with theory, we remove those observations for which the overlap assumption is not fulfilled, which make up 0.8% of the sample (=23 observations). The most important features[18] for predicting the propensity scores can be found in Appendix 5.8.7. Especially land-related features seem to play a considerable role in determining the propensity scores, which is in line with previous findings in the literature (e.g. Arata & Sckokai, 2016; Mennig & Sauer, 2020; Pufahl & Weiss, 2009). The GRF algorithm selected overall 108 features for estimating the propensity scores.

The same set of features as above was used to train the regression forest for the environmental indicators. Feature importance of the environmental outcome variables ($\hat{m}(x)$) are summarized in Appendix 5.8.8. Especially the share of grassland as well as crop and livestock outputs produced appear to be recurring important determinants of these indicators.

5.6.2 Heterogeneous Treatment Effects of AES

Estimated treatment effects seem to vary considerably across farms for all four indicators as depicted in Figure 5.3, thus indicating that the environmental effects of AES are indeed heterogeneous across farms. Table 5.3 summarizes the partici-

[18] Feature importance is defined in terms of the number of splits on a feature. For instance, if the feature importance value of a variable is 0.16, it means that the causal forest spent 16% of its splits on that variable. This measure should not be interpreted in a causal fashion, e.g., a feature with low importance is not related to propensity. This is because if two covariates are highly correlated, the trees might split on one covariate but not the other. If one was removed, however, the tree might split on the other.

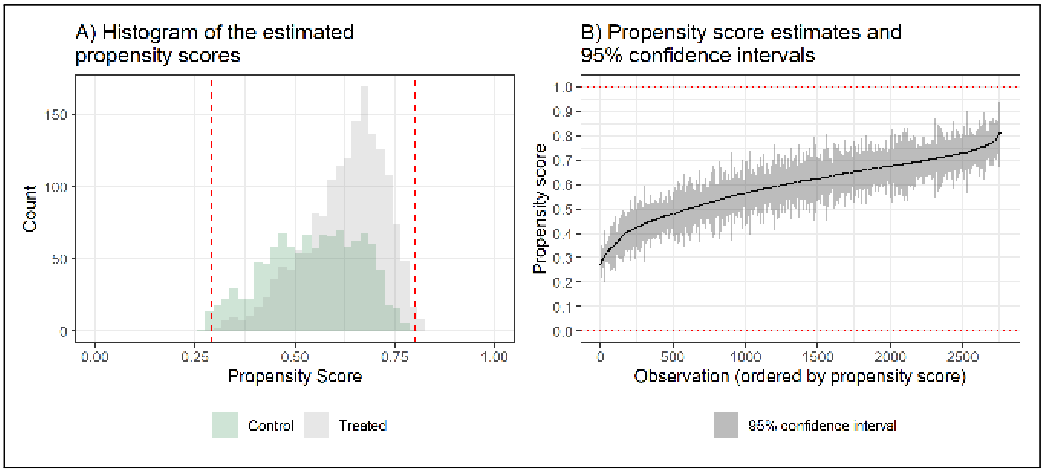


Figure 5.2: Summary of the propensity scores obtained from the step-1 propensity forest.

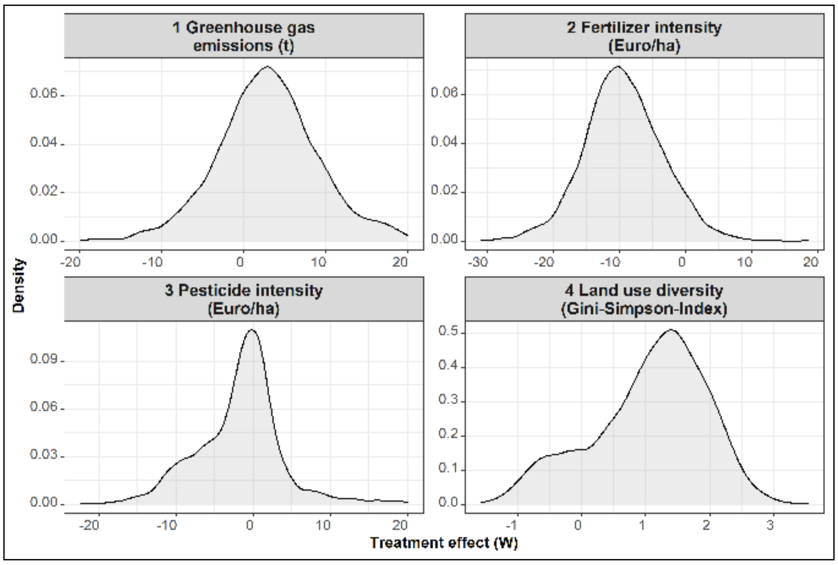


Figure 5.3: Causal forest result: Distribution of the HTE estimates for the four environmental indicators.

pation effects on the different environmental outcomes (see also App. 5.8.10 and App. 5.8.11).

As for greenhouse gas emissions, approx. 30% of the observations show the expected negative sign (Figure 5.3, upper left panel; Table 5.3). Surprisingly, a large majority of treated farms seem to have increased their emissions. Yet significant GHG effects could only be detected in 4.4% of all cases. Significant emission growth as a consequence of scheme participation on the other hand amounts to around 12 tons per farm. Expressed in terms of the average farm-level GHG emission quantity in 2014 (Table 5.1), this means an increase by 2.6%. As stated earlier, however, most farms in the sample do not show any significant treatment effect concerning GHG emissions. Different results were obtained by Dal Ferro et al. (2016), who found a slight decrease in GHG emissions as a result of AES. In light of the low GHG effects discovered in our study and the fact that the thematic coverage of AES was extended to climate objectives following the 2009 CAP Health Check and that in the current funding period AES are even referred to as 'agri-environment-climate schemes', emphasizing current and future climate change mitigation and adaptation efforts, the design of the measures needs to be reconsidered.

Table 5.3: The impact of agri-environment schemes on different environmental indicators.

	Environmental Indicator			
	GHG Emissions (t)	Fertilizer Intensity (Euro/ha)	Pesticide Intensity (Euro/ha)	Land Use Diversity (Index)
Full sample				
Mean treatment effect	3.57	-9.37	-1.41	1.06
SD treatment effect	7.86	6.02	6.44	0.89
Precentage of N with treatment effect < 0	29.4	93.7	61.7	15.0
Precentage of N with treatment effect > 0	70.6	6.3	38.3	85.0
Subsample 1 (Treatment effect $<$ 0 at 95% confidence level)				
N	6	908	183	28
Share in full sample (%)	0.2	33.2	6.7	1.0
Mean treatment effect	-10.79	-14.30	-10.28	-0.94
SD treatment effect	1.77	4.15	3.36	0.20
Subsample 2 (Treatment effect $>$ 0 at 95% confidence level)				
N	114	0	18	1511
Share in full sample (%)	4.2	-	0.7	55.3
Mean treatment effect	12.04	-	6.62	1.60
SD treatment effect	4.70	-	3.05	0.49

In terms of fertilizer expenditures per hectare (Figure 5.3, upper right panel), we find significant reduction effects in around 33% of the cases, and 94% show the expected sign, giving strong indication for a positive impact of AES (Table 5.3). The effect size varies from -31 to +18 €/hectare. Among the farms that show a significant reduction in fertilizer expenditures, we find an average effect of -€14. Given a price of 0.906 €/kg of pure nitrogen in 2014, this is equivalent to a decrease of 13 kg of pure nitrogen per ha (neglecting other fertilizers). The reduction effect we found seems to match priorities set in Bavarian agri-environmental policy. Other studies that do not consider farm heterogeneity in their assessment found more pronounced treatment effects with respect to fertilizer expenditures, e.g. Pufahl & Weiss (2009), Arata & Sckokai (2016), Uehleke et al. (2022) for the period between 2000 and 2006. With respect to pesticide intensity (Figure 5.3, bottom left panel), we find that 62% of sample farms show the expected reduction response. Out of these, however, only 6.7% are statistically significant (Table 5.3), which is indicative of the fact that AES might not have a large impact on pesticide expenditures per hectare. While Pufahl & Weiss (2009) find a significant ATT of AES on pesticide expenditure, our results are rather in line with the findings of Arata & Sckokai (2016), who do not find a significant treatment effect of AE schemes on pesticide intensity between 2003 and 2006 in Germany. The fact that our result suggests no to very little effect of environmental subsidies on pesticide expenditures per ha does not necessarily mean that they do not promote a reduction in the impact of pesticides on the environment. According to Möhring et al. (2019), quantitative pesticide indicators – such as the one used in this study – might fail to identify pesticide use patterns with the greatest risks for the environment. Finally, we find a positive effect on land use diversity for nearly all observations (Figure 5.3, bottom right panel). However, a significantly positive impact could only be found for 55% of all cases (Table 5.3). Considering a mean diversity score of approx. 66 (Table 5.1), the mean heterogeneous treatment effect of just above one appears to be very small. Likewise striking is that, spatially, regions with high uptake rates of measures aiming at diversifying crop rotations are not always identical with regions where the land-use diversity effect size is high – a situation which might indicate that the payments suffer from windfall effects (compare Sec. 5.3.2). Our results support findings on adverse participant selection and demonstrate that there is ample room to improve the schemes' efficiency. Besides revising the targeting of these subsidy payments as one way to achieve this goal (compare Section 5.6.3), the policy design of such measures could also be improved by moderating payments depending on the farmers' opportunity costs, increasing monitoring and strongly penalizing non-compliance (Gómez-Limón et al., 2019b; Latacz-Lohmann & Breustedt, 2019). Tailored payments, however, need to be accompanied by the efforts of farm advisors in order to increase uptake rates in regions where the scheme effect is shown to be high (Ferraro, 2008; Schomers & Matzdorf, 2013). Descriptively, all environmental indicators point toward heterogeneous treatment effects. To measure the impact heterogeneity statistically, we applied an omnibus test for treatment effect heterogeneity (Athey & Wager, 2019) for all four environmental outcomes (see Appendix 5.8.12). Clear evidence for treatment effect heterogeneity could be found for land-use diversity. This is not surprising since we found a rather large portion of significant effects for this indicator, while we only found a relatively small fraction of significant effects for fertilizer and pesticide intensity and GHG emissions. However, as noted by Athey & Wager (2019), that does not necessarily mean that there is no heterogeneity present in these outcomes. In fact, the finding that there are significant effects for only a small fraction of observations provides interesting insights by itself, which we would have missed if we adhered to traditional econometric techniques such as, e.g., linear regression or propensity score matching. This has also implications for legislation. The fact that an AES might be (in-)effective, on average, might induce flawed policy conclusions. For instance, an agri-environment program might be abandoned because it proved ineffective on average, although it might be effective for specific subgroups. The on average environmental ineffectiveness might as well just be the result of insufficient targeting. Hence, the ability to evaluate AES participation effects at the farm-level enables policy-makers to draw more nuanced conclusions.

Next, the locational setting of a farm often determines its farming context to a large extent, which is why we analyze the spatial heterogeneity of AES. The efficacy of AE schemes, as well as spatial scheme uptake is depicted in four maps in Figure 5.4. While panel A and B show the spatial distribution of agri-environmental payments and the share of farms participating in AES respectively, panel C and D map the portion among all observations that show the desired or undesired effect for any of the indicators selected. Certainly, such a comprehensive approach looking for any effect for different indicators ignores trade-off relations among environmental categories, however, it helps to easily detect whether an agri-environment program generally reaches environmental goals.[19] As Figure 5.4 demonstrates, this seems to be the case in most parts of Bavaria. Especially northern and western districts seem to benefit from AES in terms of environmental outcome. Districts in the Southern Alpine region and in (North)Eastern Bavaria (‘Bavarian forest’), on the other hand, where extensive forms of land use dominate, respond less strongly to AES. In some cases, the portion of observations with statistically significant adverse effects even reaches values of 30% there. Interestingly, there is a certain overlap between regions of high support and participation and regions of compara-

[19] Appendix 5.8.14 contains a complete map including disaggregated indicator-specific results.

tively low effects. This does not automatically mean that environmental payments are ineffective in these regions. Species richness for example was found to be rather high in these grassland-dominated areas and AE schemes might have a positive impact on biodiversity (Heinz et al., 2015). AES payments might in fact keep farmers from intensifying land use. However, the support-effect discrepancy can also point toward the existence of windfall effects and the potential for improved outcomes. Certain districts in Central Bavaria for instance show relatively low AES participation rates, but prominent effects. Encouraging farmers in such districts to participate in agri-environment measures might result in a higher AES cost effectiveness. Only looking descriptively at the spatial variance of AES does not reveal which contextual factors are specifically responsible for the AES treatment effect as different contextual factors are likely to be confounded. However, fair evidence-based targeting to improve environmental effectiveness requires the attribution of treatment effects variation to specific contextual factors (see next section).

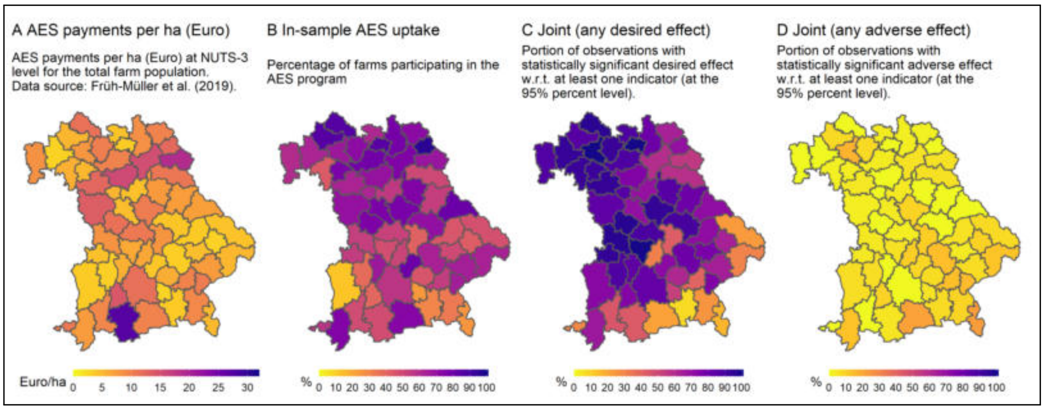


Figure 5.4: Spatial distribution (at NUTS-3-level) of A) AES payments per ha (Source: Früh-Müller et al. (2019)), B) the AES participation rate, C) percentage of observations for which any desired treatment effect w.r.t. fertilizer and pesticide intensity (€/ha), land use diversity (0-100), and greenhouse gas emissions (t) could be found, and D) percentage of observations for which any adverse treatment effect could be found.

Further, to assess the credibility of our analysis, we conducted a series of robustness tests to evaluate potential model misspecification and ommited variable bias (OVB). Appendix 5.8.15 provides a detailed summary of the robustness check results. From these tests, we can conclude that there is little evidence that our analysis suffers from model misspecification bias. However, some of the tests assessing OVB suggest that there is the possibility of bias if there exist latent confounders that are not correlated to the observed confounders. Especially if there were a lot of signal in left-out information due to unobserved confounding, our results might likely be biased. By simulating unobserved confounding using varying correlation structures, we find that, for the case of weak correlation structures, little to no bias in the treatment effect for all indicators except land use diversity. Also, especially the fertilizer intensity and land use diversity models are sensitive to stronger confounding and results become increasingly unreliable. The possibility of OVB – if we deviate from our assumption that all relevant information is latently contained in our observed data – should be taken into account when interpreting our results.

5.6.3 CATE drivers and targeting

The identification of heterogeneous treatment effects, but particularly of drivers behind these effects provides policy makers with crucial information when revising current or drafting new, targeted measures. While the practical applicability of ML in identifying HTE drivers has long been hampered by difficulties in interpreting models and their predictions, methodological advancements now allow for the identification and prioritization of features that determine outcomes.

To explain the individual farm-level treatment effect estimates, we make use of Shapley values (Shapley, 1988), a model-agnostic interpretability concept stemming from cooperative game theory, which is well-suited for complex prediction models (Lundberg & Lee, 2017; Molnar, 2019; Tiffin, 2019). Concretely, Shapley values measure the average marginal contribution of an individual variable and its values across all possible variables. For instance, a positive Shapley value of 0.8 for some feature x leads the individual prediction of the CATE to be higher than the sample mean prediction of the CATE by 0.8 units.[20] This approach allows us to assess the marginal contribution of treatment effect drivers (Tiffin, 2019) such as farm size and location, which provides additional insights as to how legislators could optimally target farms in such a way that the efficacy of AES is improved. A detailed description and further discussions on the method can be found in the Appendix 5.8.5 and Molnar (2019). We use Shapley values as suggested by Štrumbelj & Kononenko (2014) and implemented in the R package 'IML' (Molnar, 2018).

[20] In the context of heterogeneous treatment effects, the Shapley value is comparable to the interaction term effect of treatment and confounder in a in a linear regression. Appendix 5.8.5 contains a more elaborate example on the interpretation of the Shapley value.

We focused on dimensions which, according to the literature on AES, policy-makers might target to improve the efficacy of agri-environment measures. To answer the question of how these factors (yield potential, farm size, farm typology, and farm location) affect AES impact size, Figure 5.5 plots the Shapley values against the respective observed values.[21] It clearly shows that the effect size varies depending on the feature values.

The Shapley values for land use diversity with respect to yield potential, for example, suggest that the treatment effect is more prominent for farms with more favorable natural conditions (indicated by high Shapley values in relation to the sample average), which might be attributed to the higher number of land use options available to farmers in high-yield locations. Particularly striking results were found for the combinations land (i.e. farm size) and greenhouse gas emissions as well as land and pesticide intensity. In both cases, drops or jumps of Shapley values, which will be assessed in more detail below, can be observed. Larger farms participating in AES show a below-average pesticide intensity reduction effect and a lift of up to 6 tons greenhouse gases compared to the mean treatment effect of 3.57 tons. We assume that these findings are linked to the Bavarian agricultural structure where large farms (in terms of farmed land) are typically arable farms with relatively low GHG reduction potential.

Although farm typology seems to drive the effectiveness to a certain extent (Figure Figure 5.5, top right), it is the contextual dimension under investigation with the lowest impact on treatment effect size, making it least attractive to be used as targeting dimension by policy-makers.

In Figure 5.5 (bottom), we plotted Shapley values for environmental outcomes against counties, with counties on the very left of the axis being located in Central and Southeast Bavaria and districts along the axis in East, Northeast, Northwest, West and Southwest Bavaria. Taking the example of the *'Oberallgäu'* county in the Southwest of Bavaria, we find that being located in this county drives the AES effect on GHG emissions, fertilizer intensity and increases land use diversity.

To use the information coming from yield potential and farm size in the same way for targeting as farm type and location, we divided farms into groups based on their Shapley values for these categories (Figure 5.6). As cutting points serve

[21] For explorative purposes, the online supplementary material contains a graph depicting the Shapley values of the full set of contextual covariates.

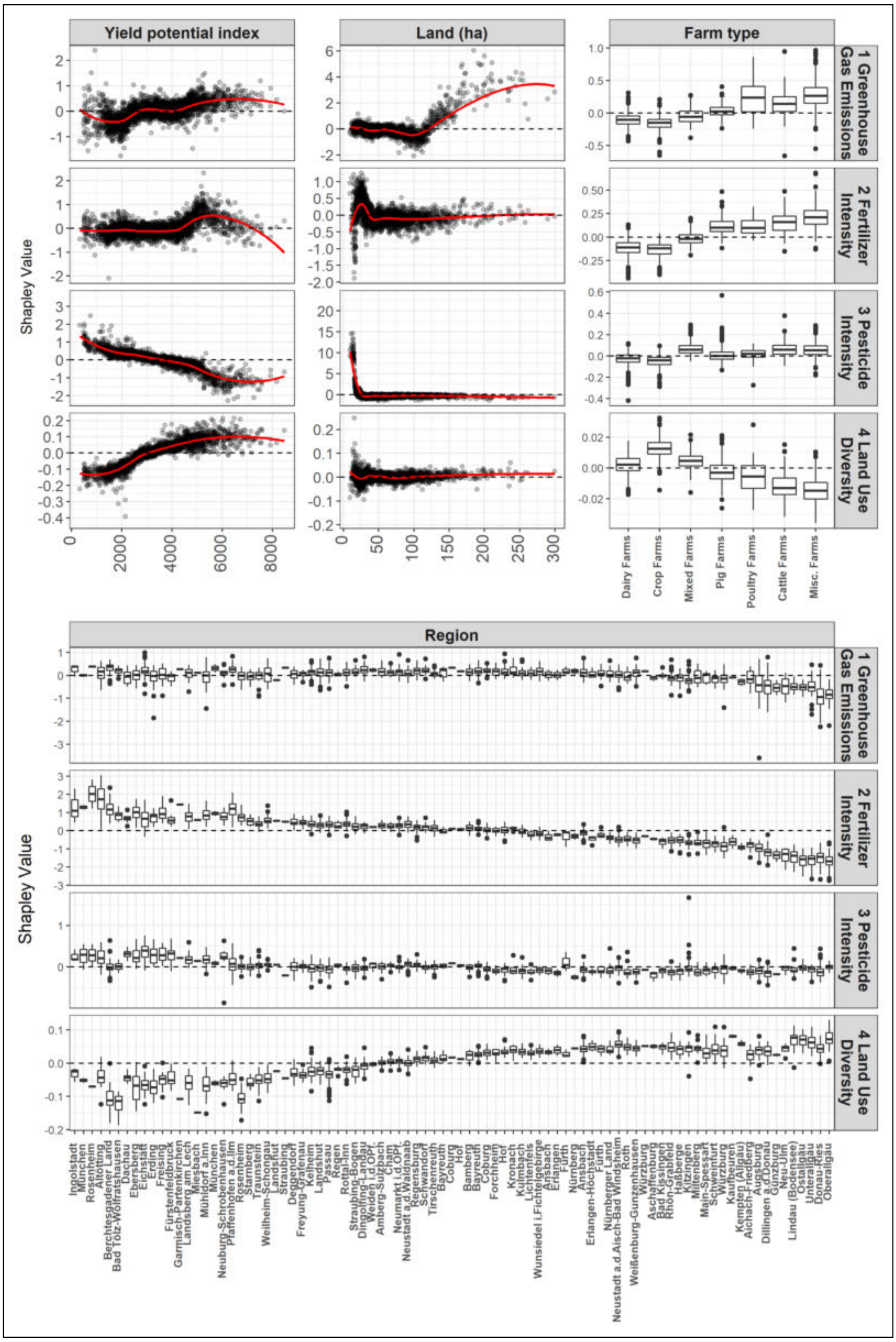


Figure 5.5: The effects of selected features on the treatment effect regarding greenhouse gas emissions, fertilizer and pesticide intensity, and land use diversity, expressed by Shapley values. They measure the average marginal contribution of an individual variable and its values across all possible variables.

the most prominent intersections of the smooth lines in Figure 5.5 with the x-axes (= zero contribution). By doing so, we are able to identify heterogeneous groups with respect to size and yield potential that mark effect size drops or jumps. E.g., in terms of pesticide intensity, for farms that are smaller than the threshold of 26 ha, the effect size is approx. 2€/ha lower (indicated by the positive Shapley value) than the mean impact as opposed to larger farms, for which the effect size is increased by 0.3€/ha. Therefore, it might be a useful strategy for legislators to target larger farms (>26 ha) if their objective is to reduce pesticide intensity. Similar patterns with varying cohort effects can be found for the other indicators and yield potential as well. Given the varying nature of these effects, it is important that policy-makers are clear about what goal they pursue when they target specific farm groups to improve the effectiveness of their measures as this could inflict negative effects regarding another goal.

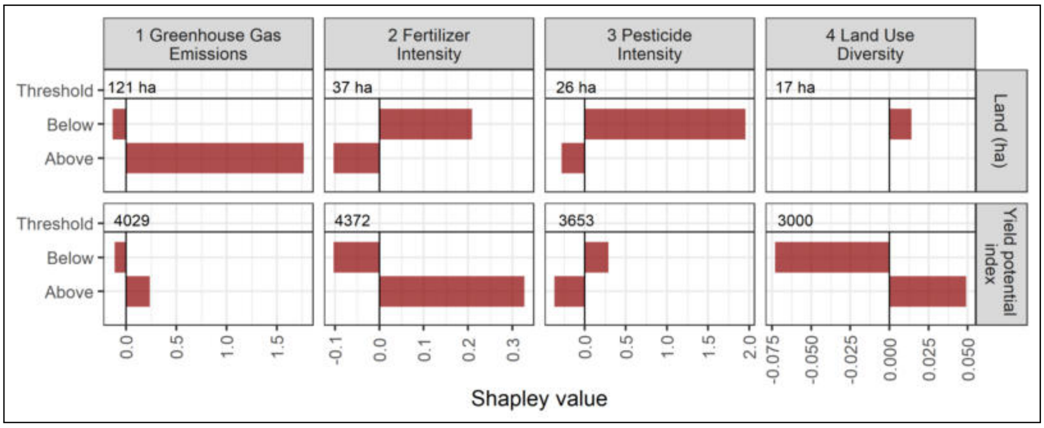


Figure 5.6: The mean effects of dividing farms into groups based on their Shapley values for land and yield potential regarding greenhouse gas emissions (t), fertilizer and pesticide intensity (€/ha), and land use diversity (0-100). The most prominent intersections of the smooth lines in Figure 5.5 with the x-axes (= zero contribution) are selected as point of division (=thresholds). This compares the mean Shapley values for the groups below and above the threshold.

Our findings on targeting ultimately describe the environmental effectiveness for specific subgroups of farmers based on the four target dimensions. Hence, this study delivers results as to *which* farms could be targeted to increase environmental effectiveness of AES. However, policy-makers might as well be interested in *how* the respective farmers can be persuaded to enrol in AES. In this context, it might be interesting to combine our results with those of, e.g., Kuhfuss et al. (2016), who suggest a collective bonus to nudge farmers into participating in AES. For instance, from Figure 5.5, we can see that being located in the "Oberallgäu" county drives up the environmental performance of farms (at least in three of the four dimensions). Legislators could consequently promote the implementation of a collective bonus explicitly for this region to nudge local farmers into participating in AES and hence increase the overall environmental effectiveness of these schemes. Other suggestions to engage farms in AES are incentive payments for their participation (Ruto & Garrod, 2009) or a reduction in transaction costs (Espinosa-Goded et al., 2013), respectively.

Finally, when interpreting these results, several important considerations should be taken into account.[22] As described in Sec. 5.3.1, there is a multitude of available agri-environment subprograms. Dichotomizing the treatment variable is invariably associated with a loss of information (Hotz et al., 2005). In an ideal situation, a policy-maker would want to learn about the heterogeneous effects for each subprogram, which would provide the largest gain in knowledge. Without the information on the farm-specific subprogram mix, it is not entirely clear if the estimated heterogeneous treatment effect is driven by effect heterogeneity (different responses to underlying multiple treatments) or treatment heterogeneity (different compositions of underlying treatments). Hence, as with other CATE studies, we cannot entirely rule out spurious discovery of heterogeneous effects (Heiler & Knaus, 2021).[23] However, if we are willing to assume that the farming context (and farm(er) characteristics) is associated with the chosen subprogram mix/AES intensity, the discussion on targeting still holds true. While we cannot test for this assumption, e.g., ART (2016) and ART (2019) suggest this might be the case. Although we cannot provide advice on the design of the programs and compare different subprograms, e.g., incentivizing farmers based on the targeting dimensions into participating in AES is still likely to improve the cost-effectiveness of AES in general without knowing the exact treatment mix. The provision of more detailed information on farms' AE (sub-)program participation might allow us to precisely disentangle effect heterogeneity and treatment heterogeneity using recent advancements in the literature on CATEs (Heiler & Knaus, 2021), which would provide additional insights.

[22] We thank an anonymous reviewer for pointing this out to us.

[23] Regardless of this fact, our approach allows us to evaluate the general environmental effectiveness of AES participation at the farm-level as described in the previous section, esp. in Table 5.3 and Figure 5.4.

5.7 Summary and concluding remarks

This paper has analyzed the environmental efficacy of agri-environment schemes in Europe in light of the post-2020 CAP debate by combining economic theory with causal forests, a novel ML algorithm based on RFs. The use of this algorithm allows to evaluate the impact of AES at the farm level and thus delivers valuable information regarding the heterogeneity of the effects of agri-environment measures. The approach presented in this study surpasses many limitations of previous attempts to evaluate the efficacy of AES based on more traditional econometric methods. Conceptually, this study is based on production theory and the potential outcomes framework.

For the empirical case of Southeast Germany, we find rather small statistically significant effects of AES on land-use diversity for approx. 55% of all observations. Regarding fertilizer expenditures per hectare, we find modest reduction effects for 30% of the sample, while we barely find any impact on pesticide expenditures. Desirable effects could be found for 7% of the sample. In terms of GHG emissions, we find mostly insignificant or adverse effects. The findings of the study point toward the direction that treatment effects of agri-environment measures on important environmental indicators have been rather small during the 2014-2020 CAP period.

Based on our results, we could explore spatial patterns of the environmental subsidy payments as well as important drivers of heterogeneous treatment effects. We found a large share of desired effects in at least one environmental dimension in almost all counties. Using Shapley values to predict the contribution of the four dimensions location, farm type, yield potential and farm size, we could confirm the hypothesis that targeting of agri-environment payments could potentially improve environmental efficacy for all environmental indicators used in this study. Targeting farms in terms of location, farm size, and yield potential by nudging for example can result in more efficient usage of environmental subsidies while targeting schemes according to different farm types does not seem to drive subsidy effectiveness. Finally, we used a battery of sensitivity tests to assess the robustness of our results in various settings.

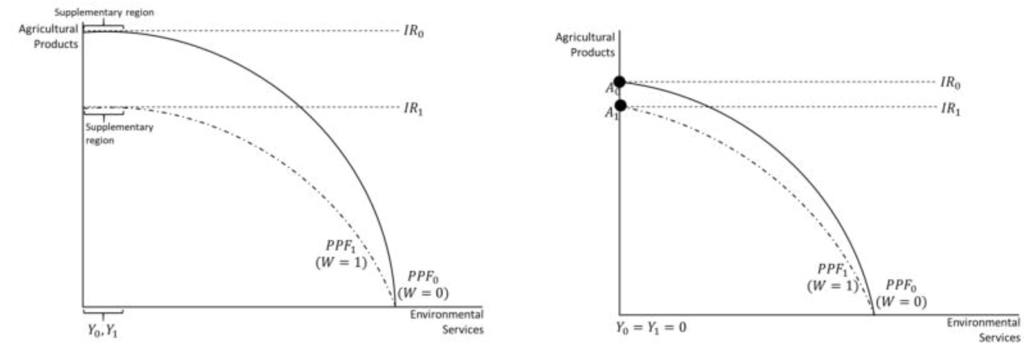
Given the novel estimation approach used in this study, there are several limitations. First, we cannot observe the effect of AES over time as we are restricted to one year in our analysis. As farms, however, must generally participate for a period of at least five years, we might miss important temporal structures as well as lagged and build-up effects of agri-environment measures. What is more, while Shapley values are useful to illustrate the drivers of impact heterogeneity, they do not account for estimation uncertainty. Introducing uncertainty to local explanations would be an important addition to the literature. Furthermore, our robustness checks indicate that there might be the possibility of unobserved confounding, which should be taken into account when interpreting the results. Next, the data do not allow for a more precise analysis of the differences across subschemes that might be targeted toward different environmental services. Also, we are limited in the choice of available environmental indicators. Except for the case of GHG, our indicators do not measure direct environmental impacts like water pollution or soil degradation. Therefore, they do not allow for a more holistic assessment of the environmental efficacy of agri-environment measures.

The findings of this study have several implications for the future of the CAP debate. First, legislators have to take into account the fact that AES have heterogeneous consequences when it comes to the environmental performance of farms. This is of particular importance when it comes to designing novel AE schemes. Second, policy-makers can potentially increase the overall environmental efficacy of AES when they improve their policy targeting such that aspects like spatiality and farm size are taken into account. Farms with high predicted participation effects could be encouraged to participate in AES through different approaches, such as paying a collective cohort bonus, reducing transaction costs, linking payments amounts to site conditions, introducing spatially-coordinated auctions for conservation contracts or other incentive payments. Third, existing AE measures appear to have very little effect or additionality in several environmental dimensions such as climate change mitigation, clean water and soil health – as approximated by our indicators. If the environmental sustainability of farms should be further improved, European legislators need to reconsider and revise existing AES.

Last, we would like to outline potential avenues for future research. One important extension to our analysis would be the assessment of subprogram-specific heterogeneous treatment effects. If there was information on specific subprograms, it might be possible to look at specific subprograms individually by controlling for the participation in other subprograms in addition to the contextual variables. Alternatively, Heiler & Knaus (2021) propose a flexible nonparametric decomposition method for the estimation and statistical inference of effect heterogeneity and treatment heterogeneity. A necessary precondition for this would be the provision of more detailed data on AES, however. It would also be interesting to see similar studies on different regions, and in different time periods, and compare the results of such studies. Furthermore, it would be insightful to include more informative environmental indicators, as they would provide a clearer picture in terms of the environmental impact of AES.

5.8 Appendices

5.8.1 Production possibilities in the case of a supplementary and competitive relationship



(a) **Supplementary case:** The farm might choose any output combination in the supplementary region.

(b) **Competitive case:** The rational farm will produce at zero environmental output.

Figure 5.7: Extension to Figure 5.1 illustrating two simple cases for (partly) supplementary (a) and competitive (b) output relationships.

5.8.2 Agri-environmental schemes in Bavaria

Note: This description can also be found in Mennig & Sauer (2020).

In line with the large variety of farming systems and landscapes in Bavaria, the federal state's 2007–2013 Rural Development Programme, which due to delays in the approval of the 2014-2020 program was still effective in 2014, included AES tailored to different agricultural subsystems. The schemes were and still are part of two programs, the Nature Conservation Programme (Vertragsnaturschutzprogramm, VNP), and the Bavarian Cultural Landscape Programme (Bayerisches Kulturlandschaftsprogramm, KULAP). The KULAP is the core funding instrument of Bavarian agri-enviromental policy, initiated as early as 1988. Individual KULAP measures of the 2007-2013 programming period were subsumed in the categories organic farming, measures for the farm segment grassland, measures for the farm segment arable land, field-specific grassland measures, field-specific measures for arable land and measures for special farming practices. In total, there were 14 individual KULAP measures, categorized according to the above-mentioned aspects. Some of the measures were further subdivided following different levels of restriction. The grassland measures as well as the measures for special farming practices mainly aimed at an extensification of production. They primarily restricted the use of mineral fertilizers and/or pesticides and set a limit to livestock units per hectare. Measures for arable land encompassed the implementation of diversified crop rotations, planting of cover crops, low-till methods, agro-environmental concepts or the conversion of arable land to grassland. Payments for special farming practices were mainly granted for measures with a focus on the conservation of the Bavarian cultural landscape. Being horizontal according to the Rural Development Programme, these measures were in fact only applicable in certain parts of Bavaria, e.g. in the alpine region or in the wine-growing districts in the north. In order to keep farmers from switching to different types of land use or abandoning farming, these measures knotted payments to environmental farming practices such as grazing or limiting the use of pesticides. Basically, all KULAP measures involved some kind of restriction, either concerning the use of certain inputs or the farmer's choice of how to cultivate the agricultural land.

The VNP, on the other hand, included very specific schemes applicable only to a small number of farms in nature conservation areas. It encompassed another 18 individual measures, some of which were further subclassified according to varying constraints, for four different biotope types: arable fields, grassland, pastures and ponds. For each biotope type, basic schemes with a rather low restriction level were offered to be combined with more ambitious schemes. Generally, the VNP schemes were more restrictive than the KULAP measures, ranging from measures with a total ban on fertilizers and pesticides on arable land as well as on grassland to leaving arable land fallow or maintaining extensive orchards on arable land, grassland or pastures. Given its eligibility constraint based on pre-defined nature conservation areas as well as its stricter guidelines, the VNP supported an agricultural area of around 65,000 hectares only, whereas KULAP payments were granted for around 1,234,000 hectares in the 2007-2013 programming period.

5.8.3 Further details on the causal forest estimator

In order to deploy the causal forest for our purpose, we define the parameters of interest as $\theta(x) = \{m(x), e(x), \tau(x)\}$ for each farm i in a random subsample of the training data. Each household has an observable quantity O_i and auxiliary features X_i; for $m(x)$ we define $O_i = \{Y_i\}$, $O_i = \{D_i\}$ for $e(x)$, and $O_i = \{Y_i, D_i\}$ for $\tau(x)$. Within this framework, Athey et al. (2019) define an optimality criterion $\partial(C_1, C_2)$ as to how an individual tree of a subsample s^{tr} splits the covariate space $\mathbf{X_i}$ of the parent node P into binary regions (C_1, C_2) to greedily[24] maximize the heterogeneity of $\hat{\theta}$ across the children nodes (C_1, C_2):

$$\partial(C_1, C_2) = \frac{N_{C_1} N_{C_2}}{N_P^2} \left(\hat{\theta}_{C_1}(s^{tr}) - \hat{\theta}_{C_2}(s^{tr})\right)^2 \tag{5.5}$$

where ${}^{N_{C_{1,2}}}/_{N_P}$ is the fraction of training examples $i : X_i \in s^{tr}$ belonging to the two children nodes C_1 and C_2 obtained from the parent P. Our parameter of interest $\theta_{C_j}(s^{tr})$ is identified by locally estimating equations of the form:

$$\mathbb{E}\left[\psi(O_i) \mid X_i = x\right] = 0 \tag{5.6}$$

where $\psi(O_i)$ is a moment condition that depends on the parameter of interest. The solution to (5.6) can be obtained through:

$$\hat{\theta}_{C_j}(s^{tr}) \in \underset{\theta, \nu}{\operatorname{argmin}} \left\{ \left\| \left(\sum_{i \in s^{tr}: X_i \in P}^{N} \psi(O_i) \right) \right\|_2 \right\}. \tag{5.7}$$

In order to find the optimal split, Eq.5.7 is solved for multiple random splits of $\mathbf{X_i}$, where the split that maximizes the optimality criterion $\partial(C_1, C_2)$ is selected (5.5).[25] The moment functions for the conditional mean estimation of $m(c)$ and $e(x)$ take on the form $\psi(Y_i) = Y_i - \theta(x)$ and $\psi(D_i) = D_i - \theta(x)$, respectively. As for the CATE model, the simple linear model $Y_i = c(x) + \tau(x)\ D_i + \epsilon$ with

[24] *Greedy* means that an optimal choice is made at each step rather than considering the entire tree when trying to find the optimal split.

[25] This procedure could now be repeated until a certain stopping criterion is reached, e.g., when a minimum size of observations per node is left. However, this would mean to estimate Eq.5.7 over and over again, which would be computationally very demanding. Therefore, an approximate criterion $\tilde{\partial}(C_1, C_2)$ based on gradient-based approximations for $\hat{\theta}_{C_j}$ is optimized, which eventually leads to computationally less expensive Breiman splits. A detailed description of this procedure can be found in Athey et al. (2019).

$c(x)$ being an intercept term serves as starting point. From this model, a doubly robust residual-on-residual linear moment function is derived, which writes

$$\psi_{\tau(x),c(x)}(\widetilde{Y_i}, \widetilde{D_i}) = \left(\widetilde{Y_i} - \beta(x) \quad \widetilde{D_i} - c(x)\right)\left(1 \quad \widetilde{D_i}^T\right)^T \tag{5.8}$$

where $\widetilde{Y_i} = Y_i - \hat{m}(x)^{oob}$ and $\widetilde{D_i} = D_i - \hat{e}(x)^{oob}$.[26] Hence, the CATE can be identified by the local moment condition: $\mathbb{E}\left[\psi_{\tau(x),c(x)}(\widetilde{Y_i}, \widetilde{D_i}) \mid X_i = x\right] = 0$. Given the local estimating equation 5.6 and the moment function (5.8), we can train a RFs-based on trees that greedily optimize for treatment effect heterogeneity (5.7), from which we can derive similarity weights $\alpha_i(x)$.[27] Analogue to (5.8) and from the simple linear model above (5.1), we can identify the CATE by $\tau(x) = Var\left[\widetilde{D_i} \mid X_i = x\right]^{-1} Cov\left[\widetilde{D_i}, \widetilde{Y_i} \mid X_i = x\right]$. Given the forest weights $\alpha_i(x)$, the locally weighted estimator $\hat{\tau}(x)$ is

$$\hat{\tau}(x) = \frac{\sum_{i=1}^{n} \alpha(x)\left(\widetilde{D}_i - \bar{D}_\alpha\right)\left(\widetilde{Y}_i - \bar{Y}_\alpha\right)}{\sum_{i=1}^{n} \alpha(x)\left(\widetilde{D}_i - \bar{D}_\alpha\right)^2} \tag{5.10}$$

where $\bar{D}_\alpha = \sum_{i=1}^{n} \alpha(x)\,\widetilde{D}_i$ and $\bar{Y}_\alpha = \sum_{i=1}^{n} \alpha(x)\,\widetilde{Y}_i$.

[26] *oob* denotes out-of-bag predictions, i.e. these predictions are generated by using only the portion of trees that do not have that data point in the respective subsample used to generate the predictions.

[27] These weights are defined as data-adaptive kernels measuring how often the i-th farm falls into the same leaf as a test point x:

$$a_i(x) = \frac{1}{B}\sum_{b=1}^{B} \frac{\mathbf{1}\left(\{X_i \in L_b(x), i \in S_b\}\right)}{\mid \{i : X_i \in L_b(x), i \in S_b\}} \tag{5.9}$$

where $L_b(x)$ is the leaf of the b-th tree that contains the test point x and S_b denotes the subsample used to grow the b-th tree.

5.8.4 Graphical illustrations of identification

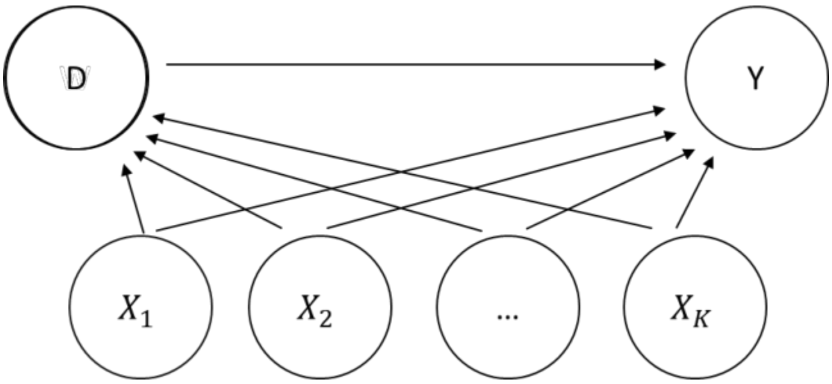


Figure 5.8: Directed acyclic graph (DAG) without unobserved confounders, i.e. the unconfoundedness assumption is fulfilled. The effect of the treatment variable D (i.e. participation in AES) on an outcome Y (i.e. environmental indicator) is identified if our model controls for all observed confounders X_1 through X_K (i.e. contextual variables) and hence all backdoor paths are closed. Connections among confounders are not shown for brevity.

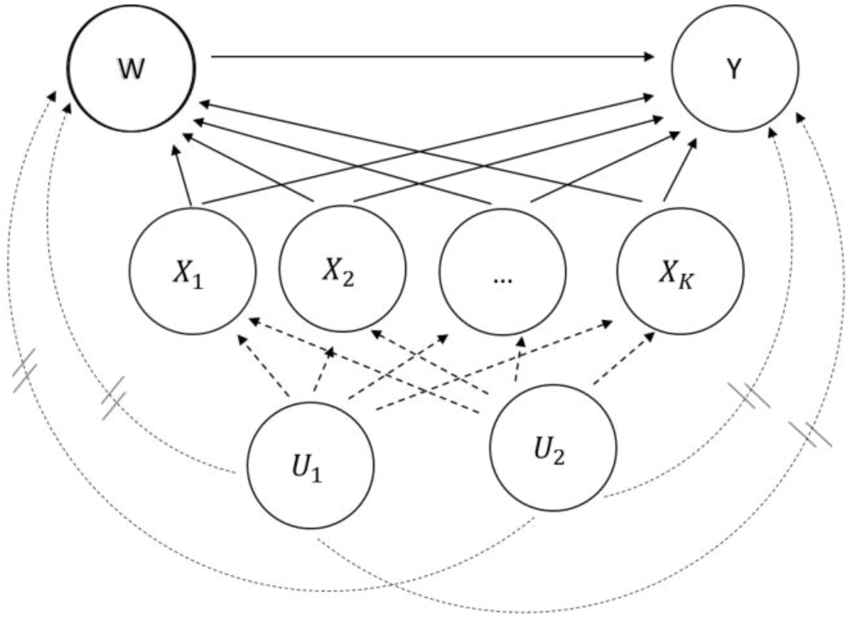


Figure 5.9: Directed acyclic graph (DAG) in an exemplary situation where two unobserved confounders (U_1, U_2) are present. The effect of the treatment variable D (i.e. participation in AES) on an outcome Y (i.e. environmental indicator) is not correctly identified if our model does not control for all observed confounders X_1 through X_K and unobserved confounders (U_1, U_2). Since U_1 and U_2 are not observable, there is no way to directly control for these confounders. Yet, under the assumption that observed and unobserved confounders are associated (arrows from U to X) and the unobserved confounders are reflected in the complex, nonlinear, and high-dimensional combination of the large number of observed confounders (latent confounder space), it might be possible to capture (most of) the variation coming from the unobserved confounder space, if the causal forest maps that latent confounder space accurately. If this is the case, the backdoor path $D \leftarrow U_{1,2} \rightarrow Y$ can be closed and the treatment effect from AES participation on environmental performance is identified. Connections among confounders are not shown for brevity.

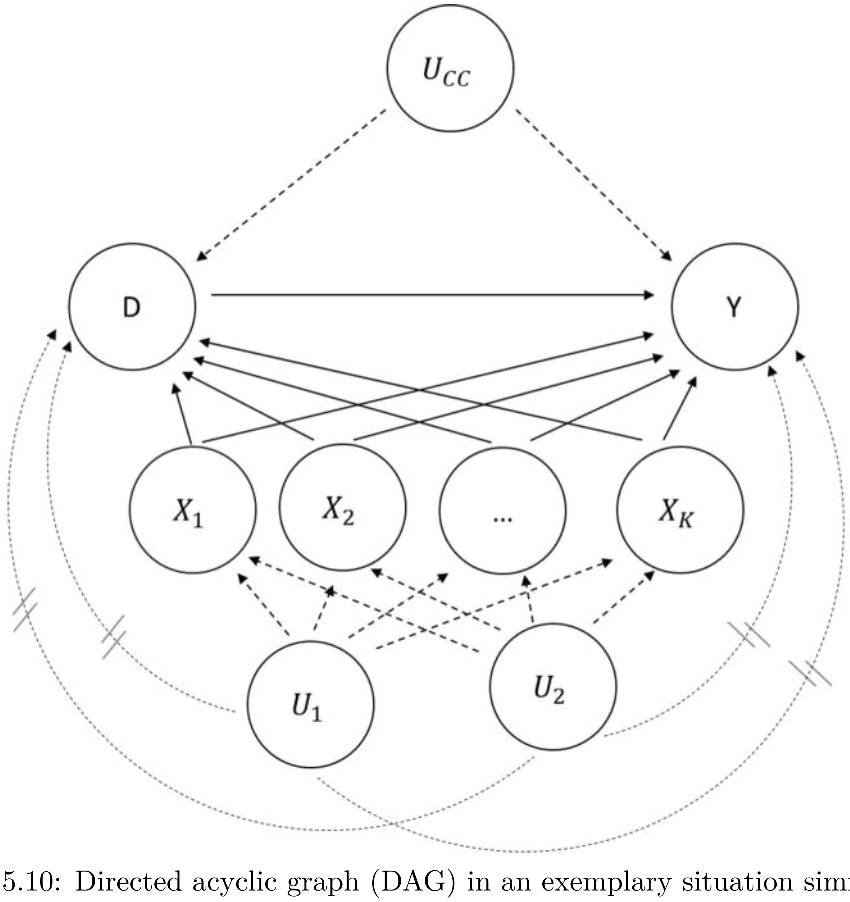


Figure 5.10: Directed acyclic graph (DAG) in an exemplary situation similar to Figure 5.9. However, there is another unobserved confounder (U_{CC}), which is not associated with the contextual variables X. In such a situation, U_{CC} cannot be represented by the set of observed variables and the treatment effect from AES participation on environmental performance is not identified and might be biased.

5.8.5 Interpretable machine learning: Shapley values

A common problem with ML prediction models is the fact that their predictive power and ability to estimate complex models comes at the expense of interpretability (Molnar, 2019). A set of model-specific as well as model-agnostic interpretability methods has been developed in recent years. Model-agnostic interpretability means that interpretation is separated from estimating/learning a model. This makes the concept very flexible in that it is not bound to any specific ML model (Ribeiro et al., 2016), i.e. it does not interfere with model particularities of the causal forest such as the honesty condition.

One concept that has been enjoying increasing attention are Shapley values. It is the only interpretability concept with a solid theory, which fulfills the axioms of efficiency, symmetry, dummy and additivity. The Shapley value is based on cooperative game theory, which fairly distributes the marginal contribution to the payoff of a game among individual players. The same context applies if one wants to interpret ML models, only that the payoff is the difference between the average prediction and the individual prediction of a test point and the features are the players. Hence, Shapley values capture the contribution of each feature (x) to the difference between the actual farm-level estimation and the sample mean estimation. Hence, Shapley values reflect each feature's relative contribution to the predicted outcome and can be seen as a special case of a marginal impact assessment, where interactions and redundancies between features are taken into account (Štrumbelj & Kononenko, 2014). Therefore, they provide an explanation why a heterogeneous causal effect model, e.g. a causal forest generate larger or smaller effect values for particular segments of the observations (Battocchi et al., 2019).

In the context of this study, we predict the heterogeneous impact of the participation in AES schemes on environmental outcome and want to study how contextual features drive this prediction. Shapley values allow us to understand the model behavior both locally and globally.

To gain a better intuition of how Shapley values work, we use a simplified example. Assume we trained a causal forest to predict the effect of participation in AES schemes on fertilizer intensity reduction (reduction in fertilizer expenditures per ha expressed in)€ contitional on three contextual variables: Yield potential, land rental price, and farmed land. The Shapley value is defined as the avaerage marginal contribution of a feature values across all possible coalitions. To obtain the marginal contribution for a specific data point and a specific feature, we need to build all possible coalitions of features as depicted for our simple case in Figure 5.11. Assume, we want to predict the marginal contribution of yield potential on the AES effect (top left panel). In the case of three features, there are four possible feature coalitions. To obtain the Shapley value, the treatment effect is estimated using coalitions 1–4 without (inserting a random number) and with the yield potential feature. Then, for each coalition the difference between the prediction with yield potential as predictor and without is calculated. The (weighted) mean of all possible coalition differences is the Shapley value for a specific observation and a specific feature. Figure 5.11 also depicts the cases for farmed area and land rental price.

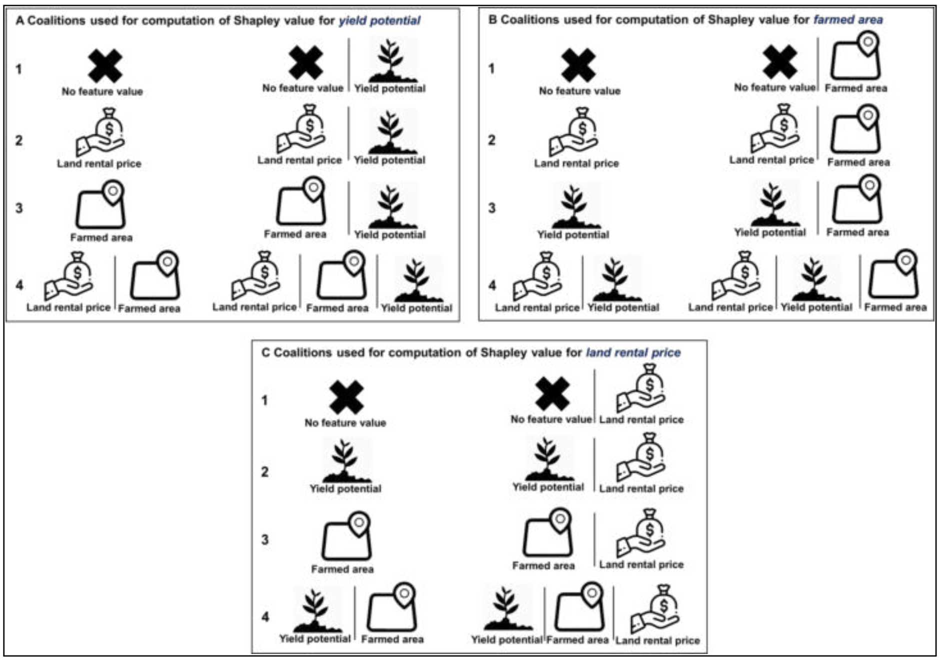


Figure 5.11: Illustration of Shapley value coalition concept by means of three contextual covariates.

Using this game theoretic approach, one can find local explanations for the model behavior at any data point, where "the payoff is the actual prediction for a particular instance less the average prediction for the entire dataset" (Tiffin, 2019). Figure 5.12 illustrates the functioning for two fictitious example farms in a sample where the average treatment effect prediction is a reduction of 35 €/ha in fertilizer expenditures. We can see that yield potential (for the farm-specific value of 6,000 points) pushes the expected outcome to the right, i.e. has a positive contribution to or impact on the prediction of the treatment effect. Similarly the farmed area (90 ha) and the land rental price (200€/ha) pushes the expected outcome to the left, hence they have a negative impact on the expected treatment effect. Finally, we end up at the farm-specific treatment effect of 10 €/ha for farm A. The sum of the Shapley values ($10 - 20 - 15 = -25$) is the difference between average prediction and instance specific prediction. Shapley values are tailored to each specific data point, i.e. although two instances have the same feature value, the Shapley value might differ. For example, Farm B in Figure 5.12 has the same yield potential as Farm A but its contribution to the expected treatment effect is twice as high. We can also see that the "payoff", i.e. the sum of Shapley values is positive for Farm B.

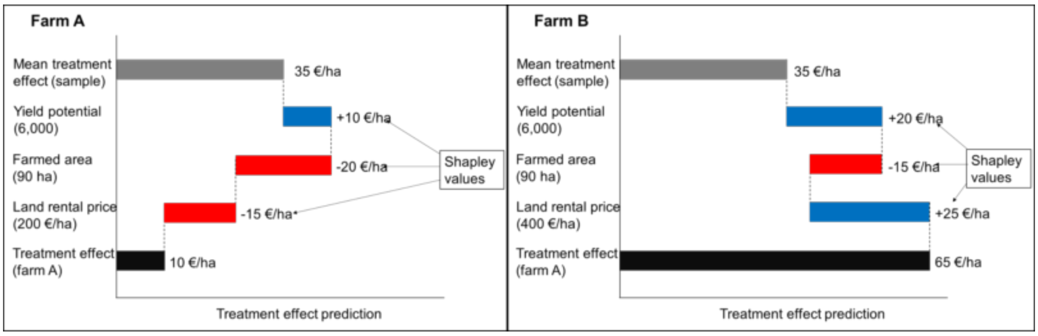


Figure 5.12: Illustration of local Shapley values expressing marginal contributions to the mean treatment effect prediction.

While it is interesting to find local impacts on specific instances, most of the time we are more interested in global explanations. To find general patterns, Shapley values can be calculated for all instances and features in the dataset. E.g., the left panel of Figure 5.13 summarizes the Shapley values of all features and contrasts them with the respective feature values (i.e. the color of the graph). For instance, in our stylized example, high land rental prices contribute positively to the treatment effect while a reversed pattern could be found for total farmed area. Taking the example of yield potential, we can obtain a more nuanced picture of the model behavior using a dependence plot (right panel of Figure 5.13). In this scatter plot, we can find a convex relationship between feature value and marginal impact of yield potential.

Finally, Shapley values also allow for the evaluation of model-agnostic relative feature importances, i.e. the overall importance of a feature for the model prediction (Figure 5.14). For that purpose, one can take the mean of the absolute values of the Shapley values. The higher this value is, the more important a feature is for the model.

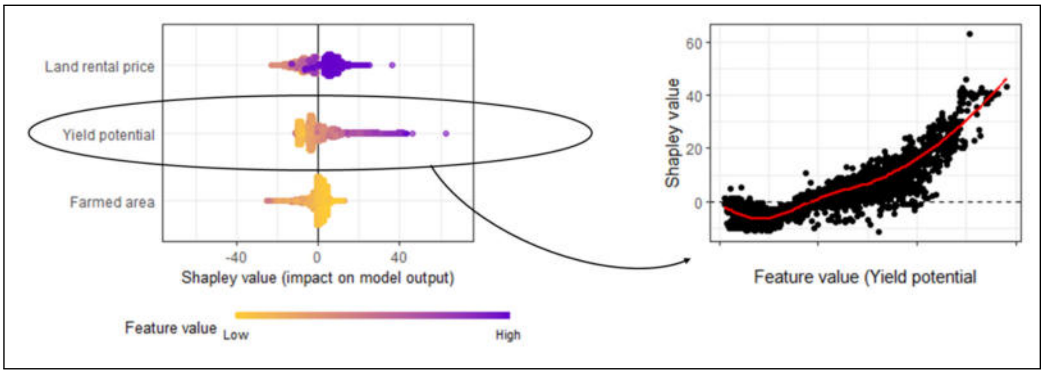


Figure 5.13: Illustration of how local Shapley values can be used for global interpretations.

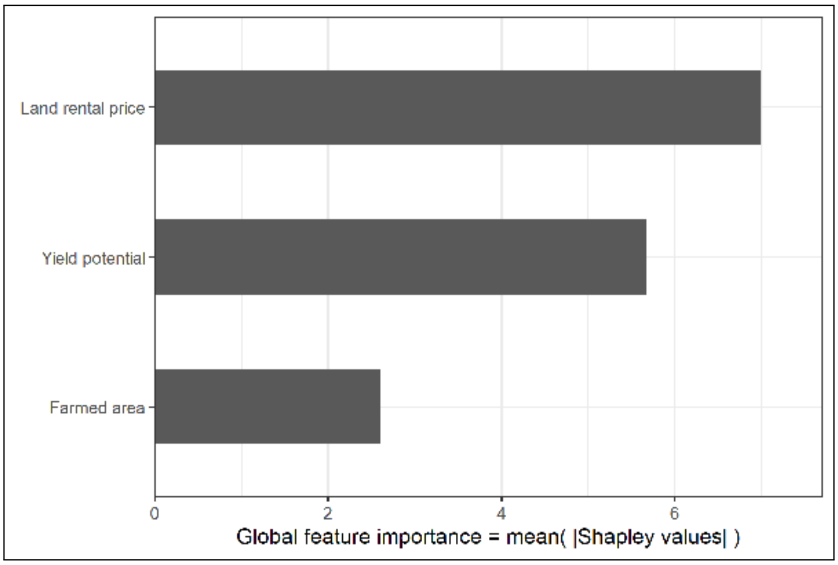


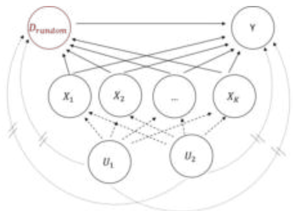
Figure 5.14: Shapley values as measure for global feature importance.

5.8.6 Sensitivity and robustness checks

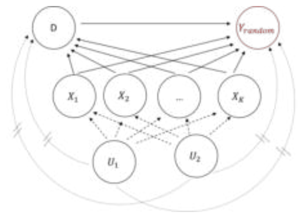
To validate our results against model misspecification and unobserved heterogeneity bias, we suggest a range of robustness checks.A visually supported description of these robustness checks against critical model assumptions can be found in Figure 5.15. First, to check for model misspecification, we replace the treatment variable D by a placebo treatment $D_{placebo}$. Second, we introduce a placebo outcome $Y_{placebo}$ instead of the observed outcomes Y. If our model is correctly specified, we expect negligible to none association between treatment and outcome in the placebo tests. Third, we add a random confounder (X_{random}) to our model and compare against our baseline model and assess the stability of our baseline estimates, which we do not expect to change.

Fourth, to explicitly analyze the model behavior regarding unobserved heterogeneity bias, we intentionally leave out the most important feature. We leave out the (three) most important confounder(s) and re-estimate the model. If the (non-linear) correlation structure of the other observed confounders properly reflect the left-out variable(s), we expect no change of the model against our baseline model. This could point towards the stability of our model against unobserved confounders that are not included in the model if they are associated with the set of observed confounders. Fifth, we leave out various observed confounders on a more systematic basis. We use principal component analysis (PCA) and the resulting loadings to detect systematic groups of confounders and leave these out, such that the combination structure of the other observed confounders are less capable to compensate for the exclusion of these variables. The retrieved model behavior could indicate how strongly it reacts to potentially completely left out confounders that are not buffered away by the combination of observed covariates.[28] Sixth, we introduce a target common cause, which is correlated with both treatment and outcome, simulating various different correlation structures between U_{target}, D and Y, which should realistically reflect a potential relationship between U_{target}, D and Y. This robustness test can be seen as indication for how sensitive the model is to the possibility of any left-out hidden factors, which might not be captured in our (latent) feature space.

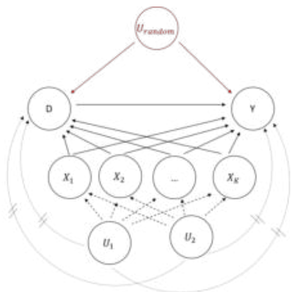
[28]Such a situation could appear if many observed features were a measure of the same latent confounder, while other latent confounders would not be covered by the observed features.



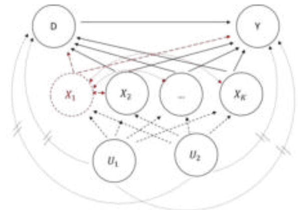
(a) The original treatment variable (D) is replaced by a random placebo treatment variable (D_{random}). If the model is correctly specified, we expect no effect of the treatment on the environmental outcome.



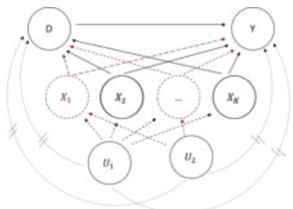
(b) The original outcome variable (Y) is replaced by a random placebo outcome variable (Y_{random}). If the model is correctly specified, we expect no effect of the treatment on the environmental outcome.



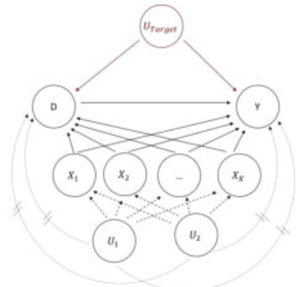
(c) We add an extra random variable (U_{random}) as potential confounder to our model. As this confounder is random, we expect no effect of the treatment on the environmental outcome.



(d) We leave out the (three) most important confounder(s) and re-estimate the model. If the (nonlinear) correlation structure of the other observed confounders properly reflect the left-out variable(s), we expect no change of the model against our baseline model, which includes all observed confounders. This could point towards the stability of our model against unobserved confounders that are not included in the model if they are associated with the set of observed confounders.



(e) Similar approach to (d). However, this time, we leave out various observed confounders more systematically. We use principal component analysis and the resulting loadings to detect systematic groups of confounders and leave these out, such that the combination structure of the other observed confounders are less capable to compensate for the exclusion of these variables. The retrieved model behavior could indicate how strongly it reacts to potentially completely left out confounders that are not buffered away by the combination of observed covariates.



(f) In this robustness check, we simulate a completely left out unobserved covariate (U_{target}) simulating various different correlation structures between U_{target}, D and Y. This gives indication as to how strong the omitted variable bias in our model could be given different correlation structures between U_{target}, D and Y.

Figure 5.15: Stylized cases reflecting the potential impact of AES participation under heterogeneous production possibilities with one agricultural and one positive environmental output.

5.8.7 Propensity forest – most important features

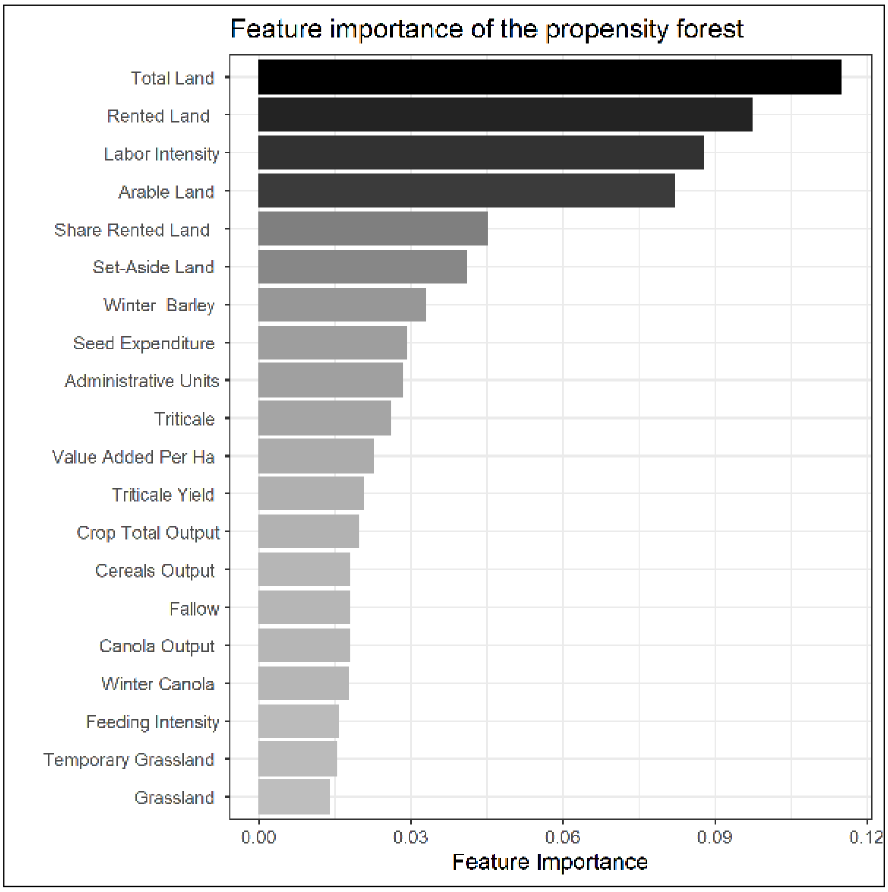


Figure 5.16: Variable importance: Depiction of the 20 most important features for the propensity forest.

5.8.8 Outcome forests – most important features

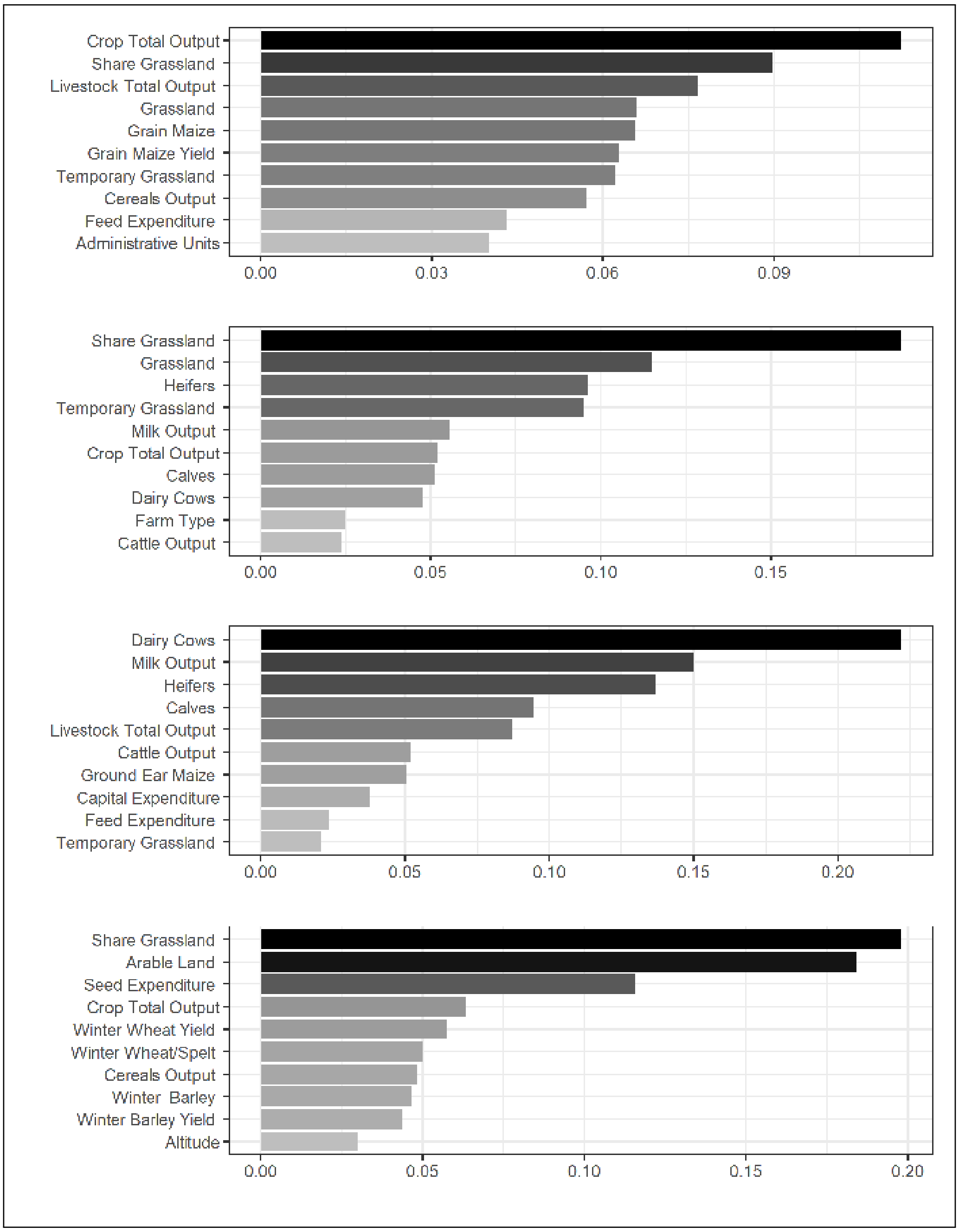


Figure 5.17: Variable importance: Depiction of the 10 most important features for each outcome forest.

5.8.9 Causal forest illustration: GHG tree

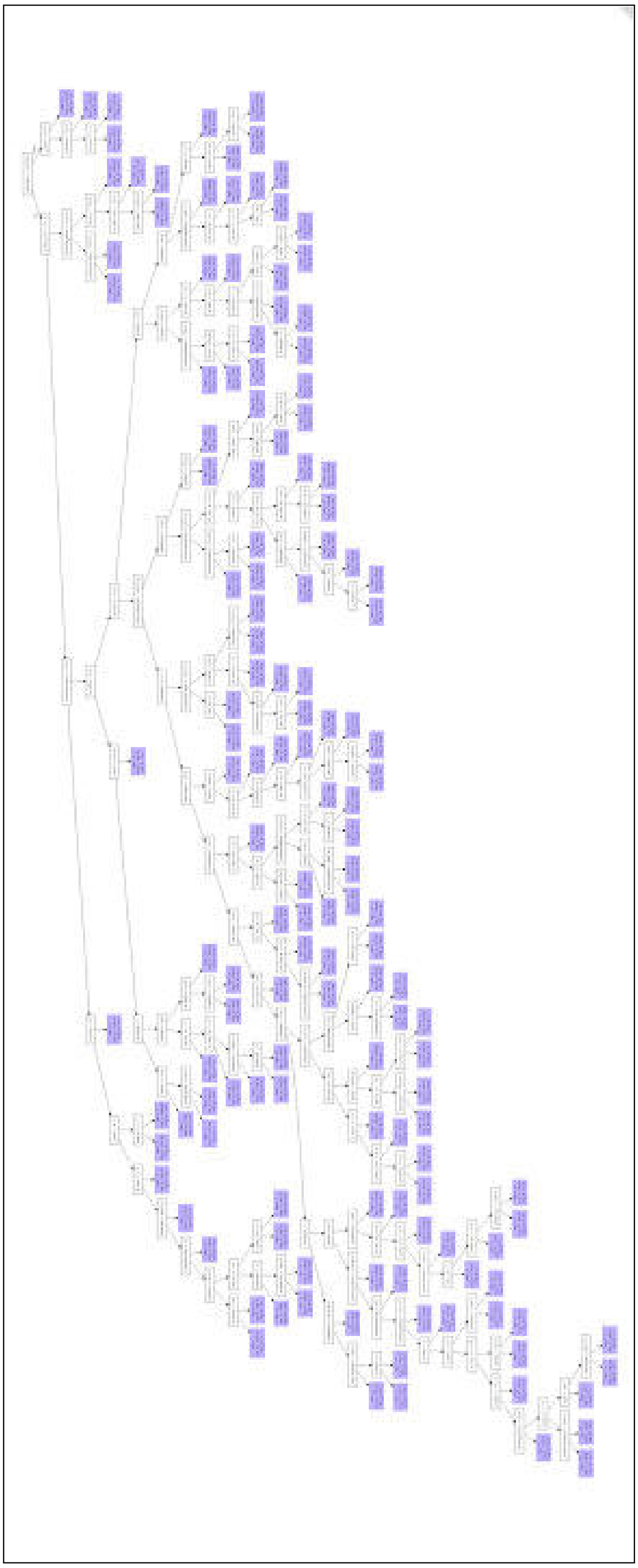


Figure 5.18: This illustration depicts a randomly selected tree from a total of 100,000 trees used to estimate the GHG causal forest.

5.8.10 Causal forest estimates and corresponding 95% confidence intervals

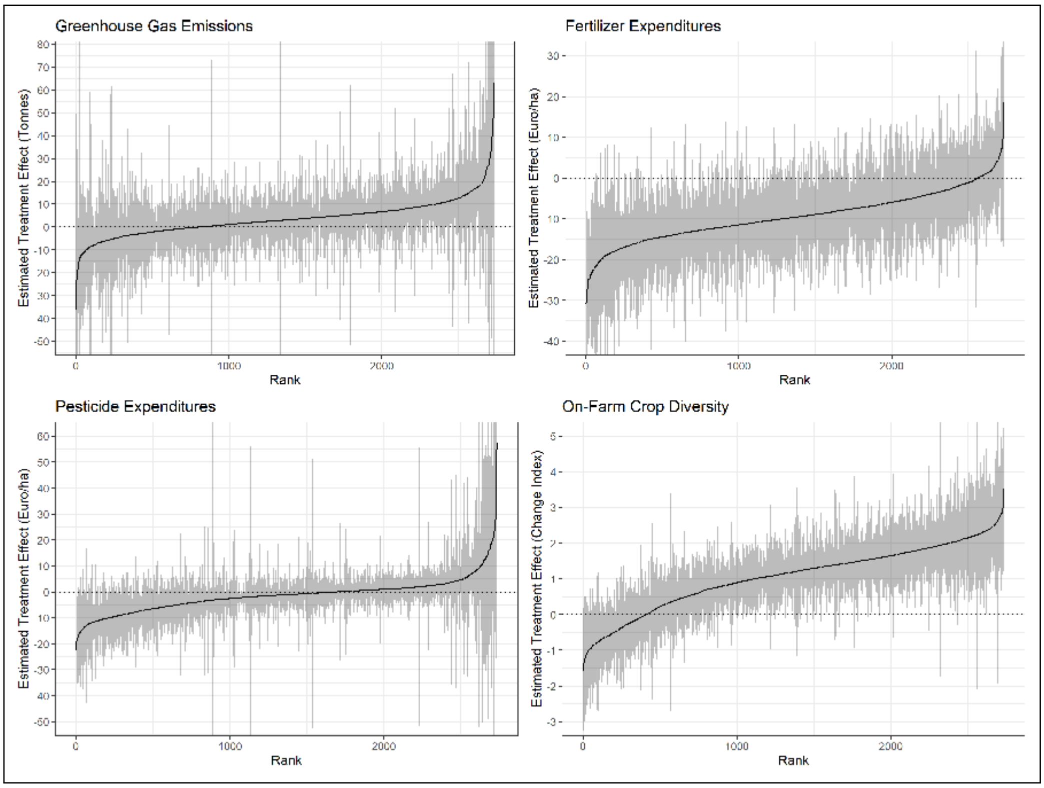


Figure 5.19: Causal forest estimates and corresponding 95% confidence intervals reflecting estimation uncertainty.

5.8.11 Treatment effects weighted by their inverse standard deviations

Table 5.4: The impact of agri-environment schemes on different environmental indicators weighted by their inverse standard deviations.

	Environmental Indicator			
	GHG Emissions (t)	Fertilizer Intensity (Euro/ha)	Pesticide Intensity (Euro/ha)	Land Use Diversity (Index)
Full sample				
Unweighted mean treatment effect	3.57	-9.37	-1.41	1.06
Weighted mean treatment effect by inverse SD	3.03	-9.4	-1.37	1.02
Subsample 1 (Treatment effect $<$ 0 at 95% confidence level)				
Unweighted mean treatment effect	-10.79	-14.3	-10.28	-0.94
Weighted mean treatment effect by inverse SD	-10.59	-13.54	-9.08	-0.9
Subsample 2 (Treatment effect $>$ 0 at 95% confidence level)				
Unweighted mean treatment effect	12.04	-	6.62	1.6
Weighted mean treatment effect by inverse SD	10.57	-	5.59	1.49

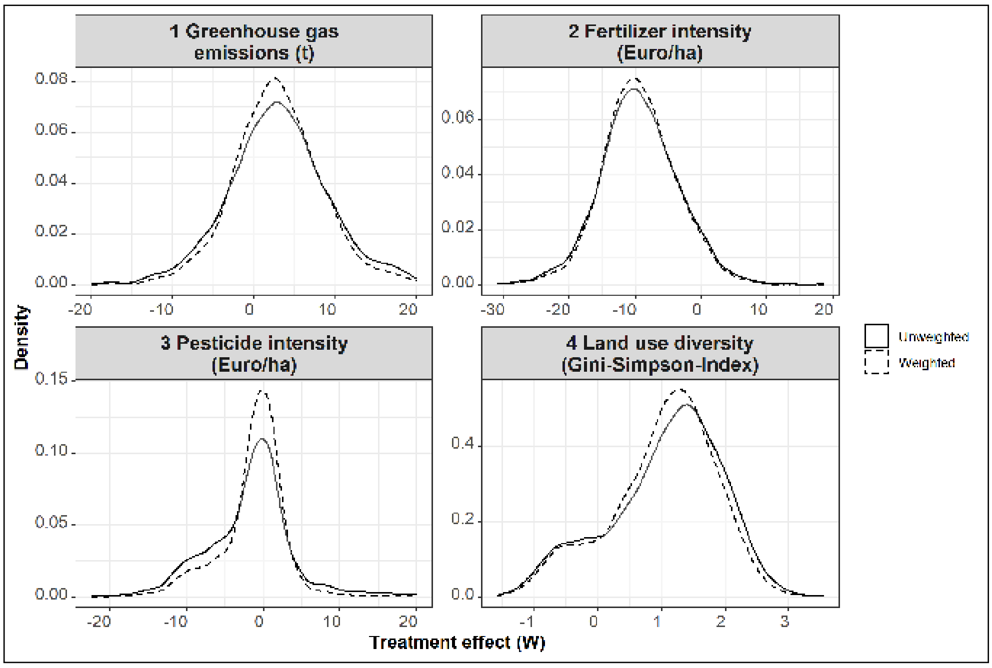


Figure 5.20: Comparison of density distributions of unweighted treatment effects (baseline) and and treatment effects weighted by their inverse standard deviations.

5.8.12 Omnibus test results for the presence of treatment effect heterogeneity

Table 5.5: Omnibus test results for the presence of heterogeneity: If the coefficient on the differential forest prediction is significantly greater than 0, then we can reject the null of no heterogeneity, which is the case for land use diversity.

	Differential forest prediction	Std. Error	t value	Pr(>t)
GHG Emission	-1.3	1.28	-1.01	0.843
Fertilizer Intensity	0.3	0.56	0.53	0.299
Pesticide Intensity	-1.2	0.93	-1.34	0.909
Land Use Diversity	0.8	0.33	2.48	0.007

5.8.13 HTE drivers: Shapley values and interaction effects

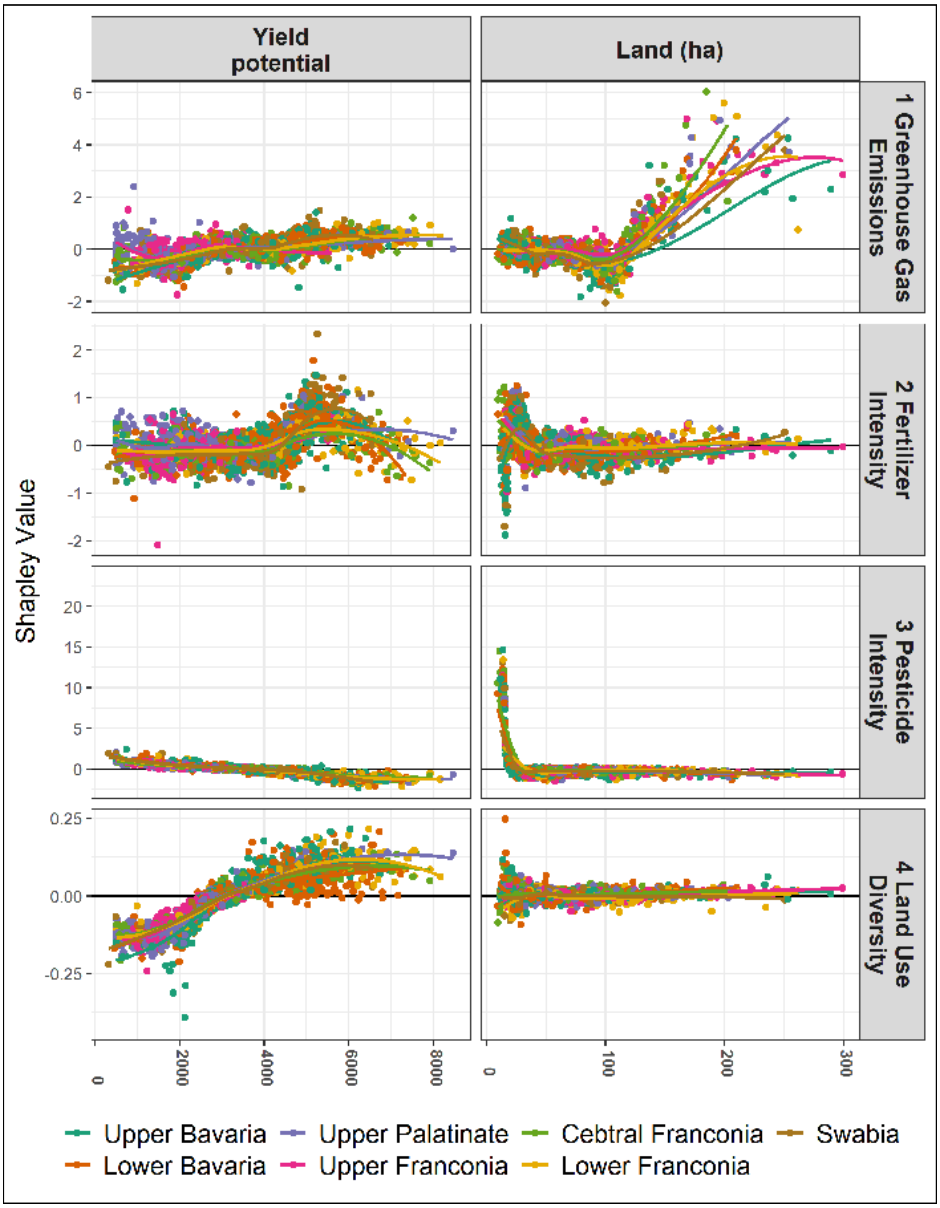


Figure 5.21: HTE drivers: Shapley values and interaction effects with administrative units ("*Regierungsbezirke*").

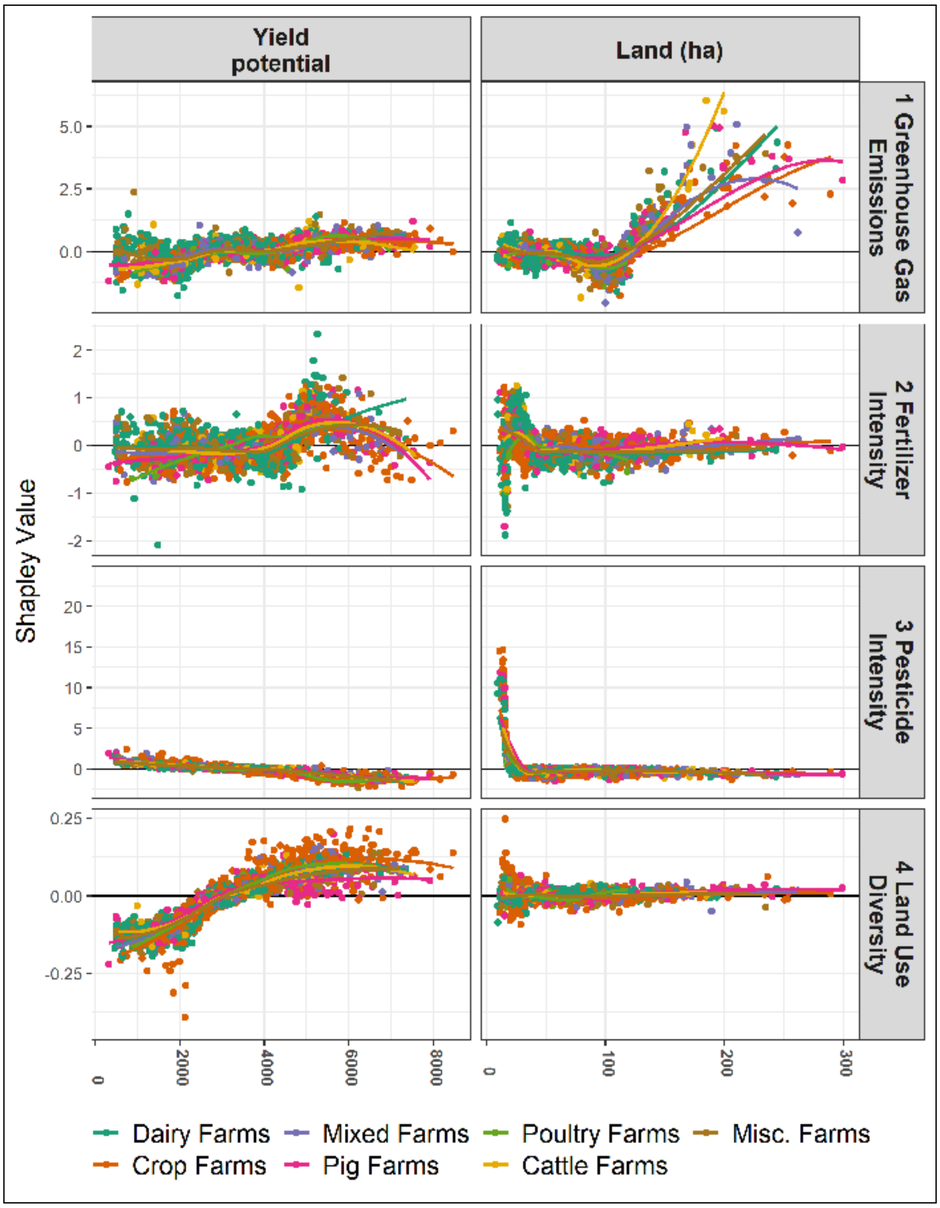


Figure 5.22: HTE drivers: Shapley values and interaction effects with farm types.

5.8.14 Spatial AES impact heterogeneity

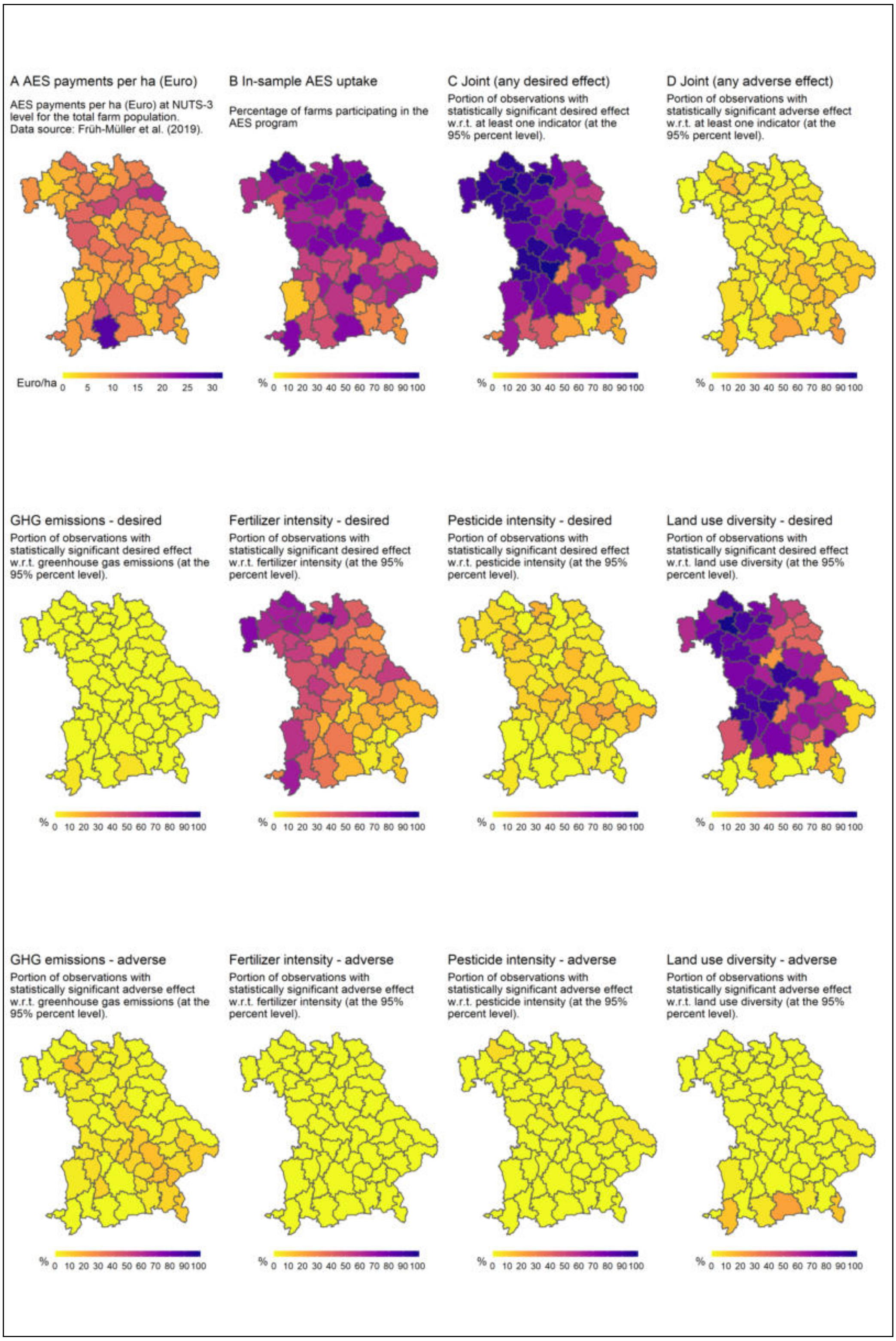


Figure 5.23: Spatial AES impact heterogeneity.

5.8.15 Robustness and sensitivity test results

One of the major concerns regarding the credibility of our results are model misspecification and omitted variable bias due to unobserved confounders. Figure 5.24 summarizes the results of these robustness checks described in Sec. 5.8.6. Regarding the model misspecification placebo tests, we observe negligible to none association between treatment and outcome for all four indicators as expected for a correctly specified model (Figure 5.24, panels A and B). Furthermore, as expected, adding another random confounder does not seem to strongly affect our estimation results either (Figure 5.24, panel C).

When we leave out the most important feature (panel D), and more conservatively the three most important features (panel E), results appear to remain almost identical. Although re-estimation seems to introduce a certain degree of statistical noise, no systematic deviation from the baseline results can be observed.[29] To further explore the sensitivity to leaving out confounders, Figure 5.24, panels F to H describe the behavior of the model when leaving out systematic groups of confounders based on principal component analysis (PCA) and their respective loadings.[30] We find that leaving out just one latent component leads to mostly weakly to moderately biased estimates. Leaving out more components leads to non-robust results. This might be due to the fact that we lose a lot of signal by leaving out too many covariates. However, by implication, this would also mean if there were a lot of signal in left-out information due to unobserved confounding, our results might likely be biased. This possibility should be taken into account when interpreting our results. Finally, to more specifically analyze the sensitivity of our results to the potential violation of the unconfoundedness assumption directly, we simulate an unobserved common cause with various correlation structures (Huber, 2020) to both AES participation and environmental performance and include this simulated confounder to our model (Figure 5.26 – Figure 5.29). For weak correlation structures, we find little to no bias in the treatment effect for all indicators except land use diversity. Moreover, especially the fertilizer intensity and land use diversity models are sensitive to stronger confounding and results

[29] However, there might be a small downward bias for pesticide intensity. Note, these result suggest that the left-out features are sufficiently correlated to the other observed confounders to be compensated for. This test does not guard against potential bias due to unobserved confounders that are not correlated to the observed confounders.

[30] A description of the left-out variables and the scree-plot of the PCA can be found in Figure 5.25. Covariates with loadings $> |0.15|$ were taken to represent a hidden component. A full list of all variables and respective loadings can be found in the supplementary materials.

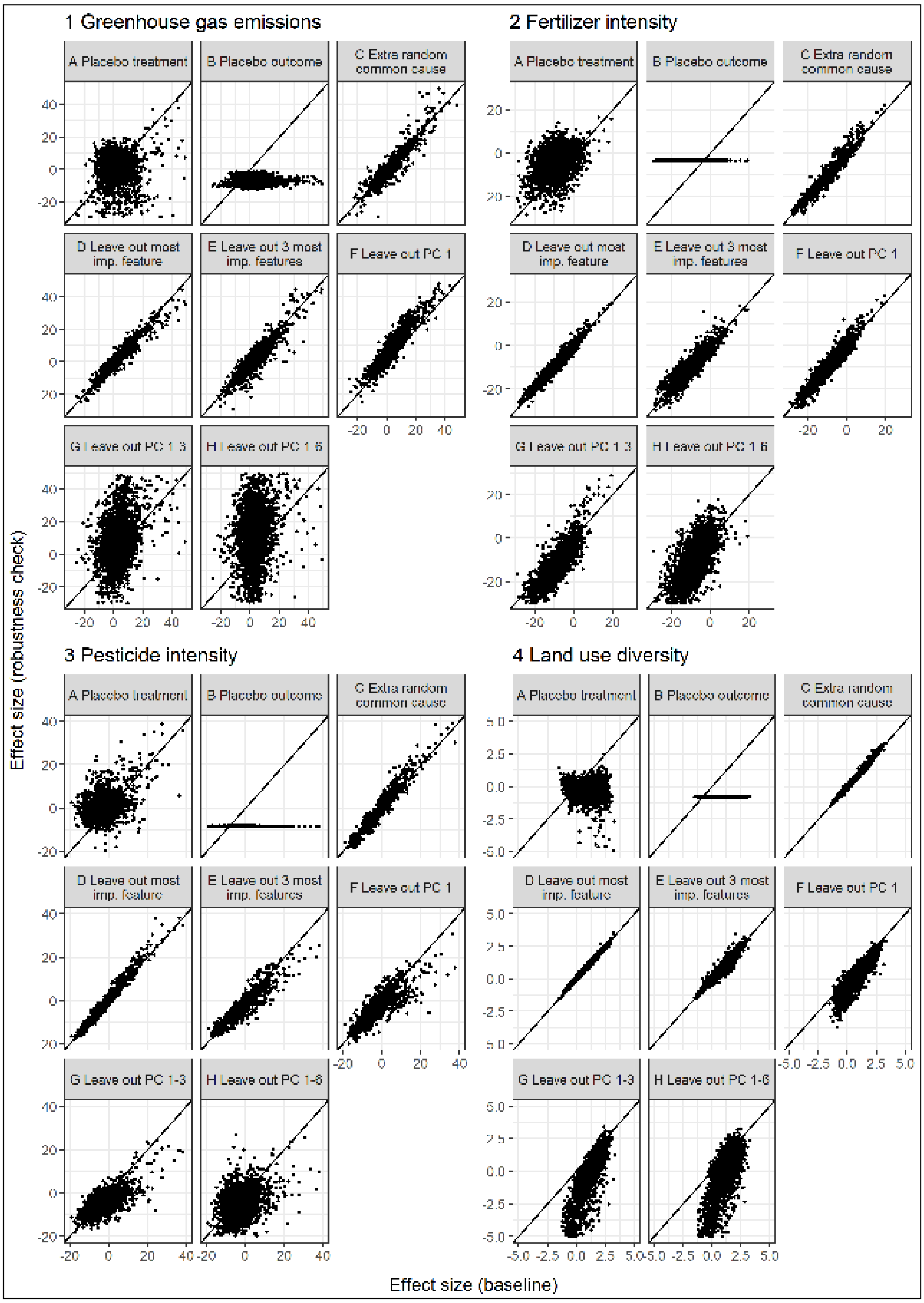
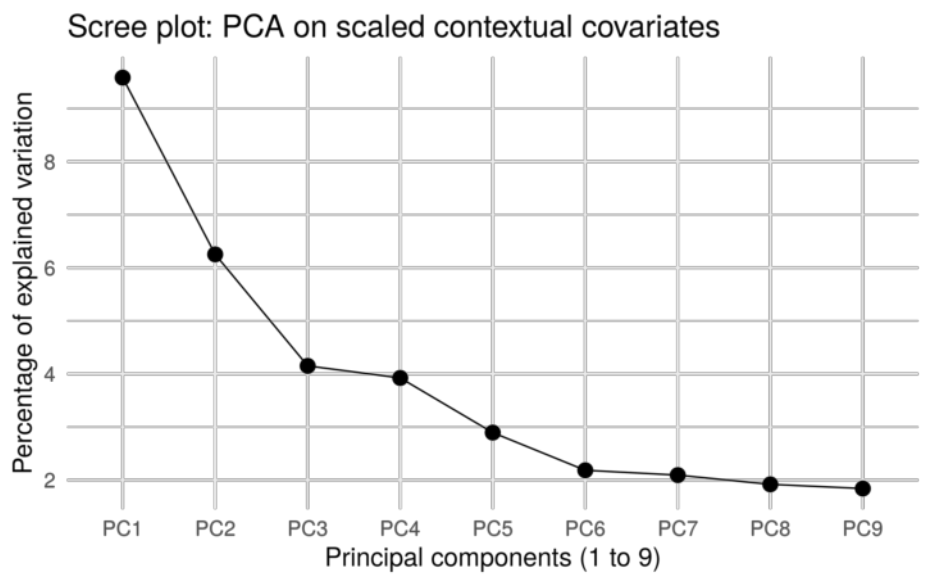


Figure 5.24: Multiple robustness checks regarding model misspecification and unobserved heterogeneity bias for greenhouse gas emissions (t), fertilizer and pesticide intensity (€/ha), and land use diversity (0-100) as described in Sec. 5.4.

become increasingly unreliable (see also Appendix 5.8.16).



Component loadings resulting from the principal component analysis on the set of contextual covariates

The variables that are left out in the corresponding robustness checks are marked in color. Only the loadings for the principal components 1 to 6 are displayed.

Principal components	Covariate groups based on component loadings > |0.15|	Robustness tests: Systematically leave out covariate groups based on PCA		
		PC 1	PC 1-3	PC 1-6
PC 1	Seed Expenditure , Share Grassland , Total Land , Rented Land , Arable Land , Whole Farm Value Added, Cereals Output , Canola Output , Crop Total Output, Winter Wheat/Spelt , Winter Barley , Winter Canola , Winter Wheat Yield			
PC 2	Family Labor , Capital Expenditure, Share Grassland , Total Land , Rented Land , Grassland , Whole Farm Value Added, Milk Output , Livestock Total Output , Dairy Cows , Calves , Heifers , Ground Ear Maize , Clover , Temporary Grassland			
PC 3	Feed Expenditure , Cereals Output , Potatoes Output , Crop Total Output, Sugarbeet Output , Pigs Output , Livestock Total Output , Weaners , Fattening Pigs , Sows , Boars , Winter Barley , Potatoes , Sugar Beet , Winter Barley Yield			
PC 4	Total On-Farm Labor, Hired Labor , Capital Intensity, Labor Intensity, Value Added Per Ha , Yield Index Unit, Whole Farm Value Added, Crop Total Output, Other Plants Output, Spring Barley , Winter Canola , Hops , Canola Yield, , Spring Wheat Yield , Triticale Yield			
PC 5	Total On-Farm Labor, Hired Labor , Labor Intensity, Yield Index Unit, Other Plants Output, Cattle Output , Male Beef , Calves , Hops , Sugarbeet Yield			
PC 6	Hired Labor , Potatoes Output , Other Plants Output, Cattle Output , Male Beef , Calves , Potatoes , Ground Ear Maize , Hops , Potato Yield, Yield Pasture, Silage Maize Yield, Yield Grass Fodder			

Figure 5.25: Description of the principal component analysis underlying the principal component (PC) robustness checks.

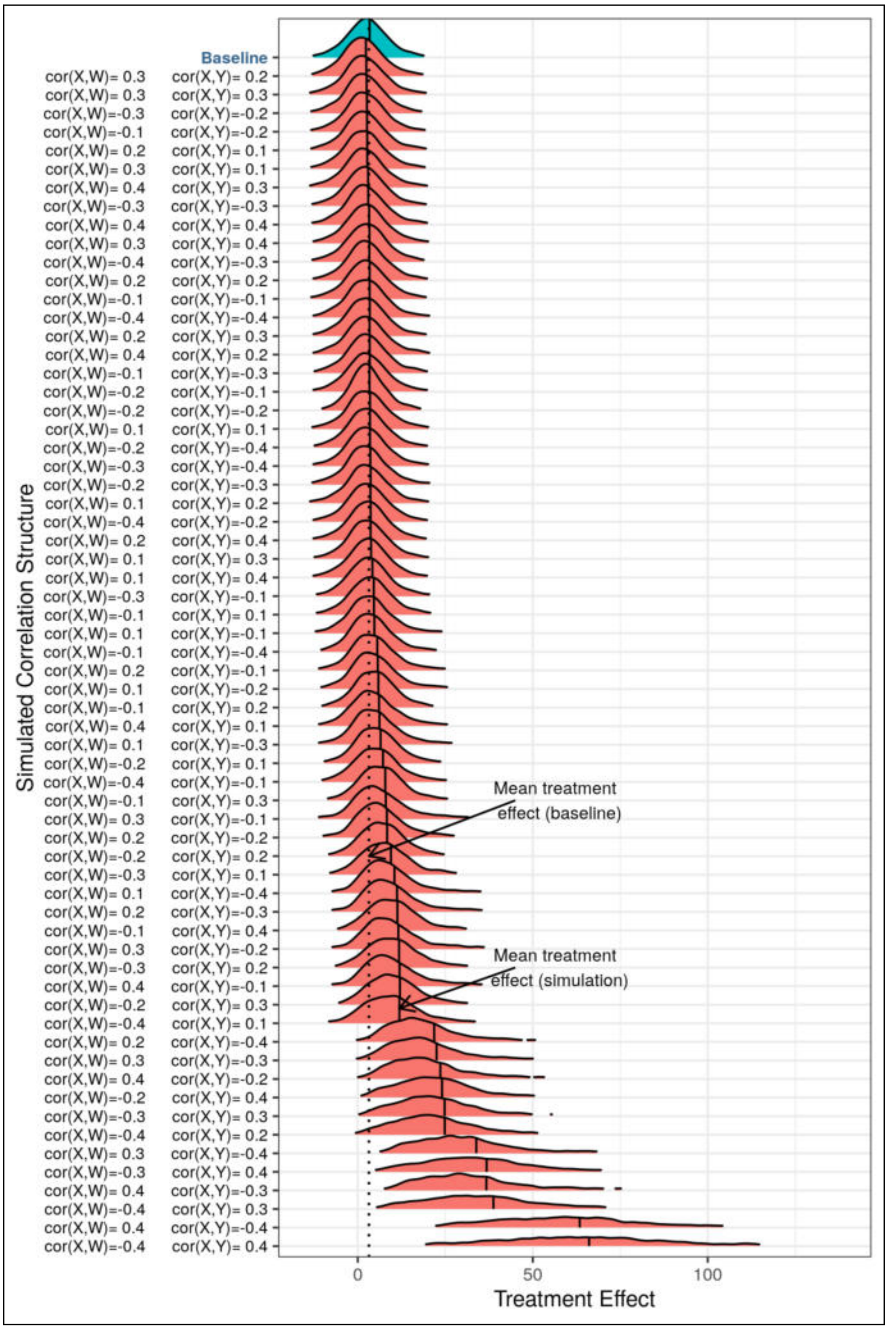


Figure 5.26: Greenhouse gas emissions: Simulation of omitted variable with different correlation structures using the approach of Huber (2020). The corresponding treatment effect distributions are depicted and compared to the baseline model.

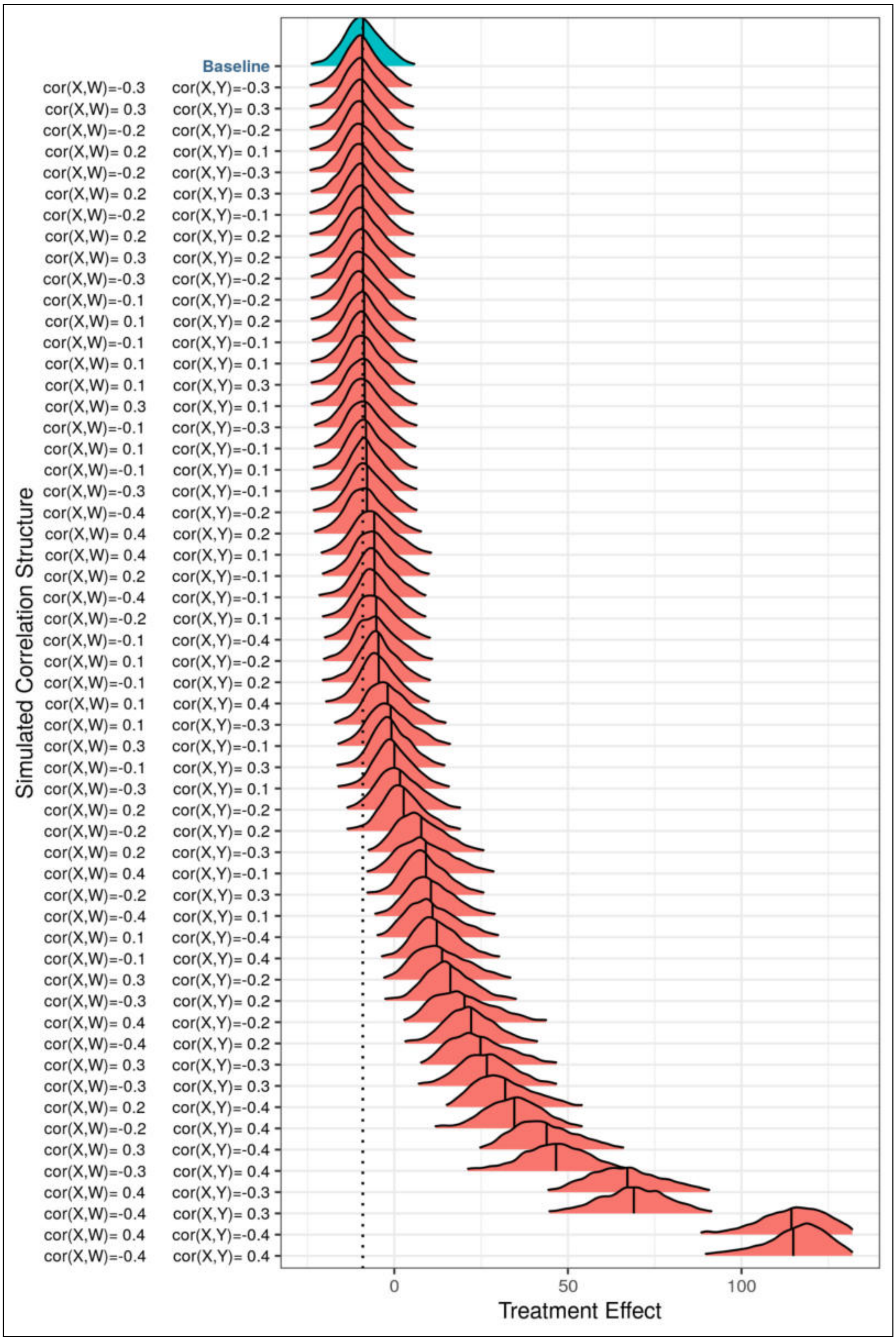


Figure 5.27: Fertilizer intensity: Simulation of omitted variable with different correlation structures using the approach of Huber (2020). The corresponding treatment effect distributions are depicted and compared to the baseline model.

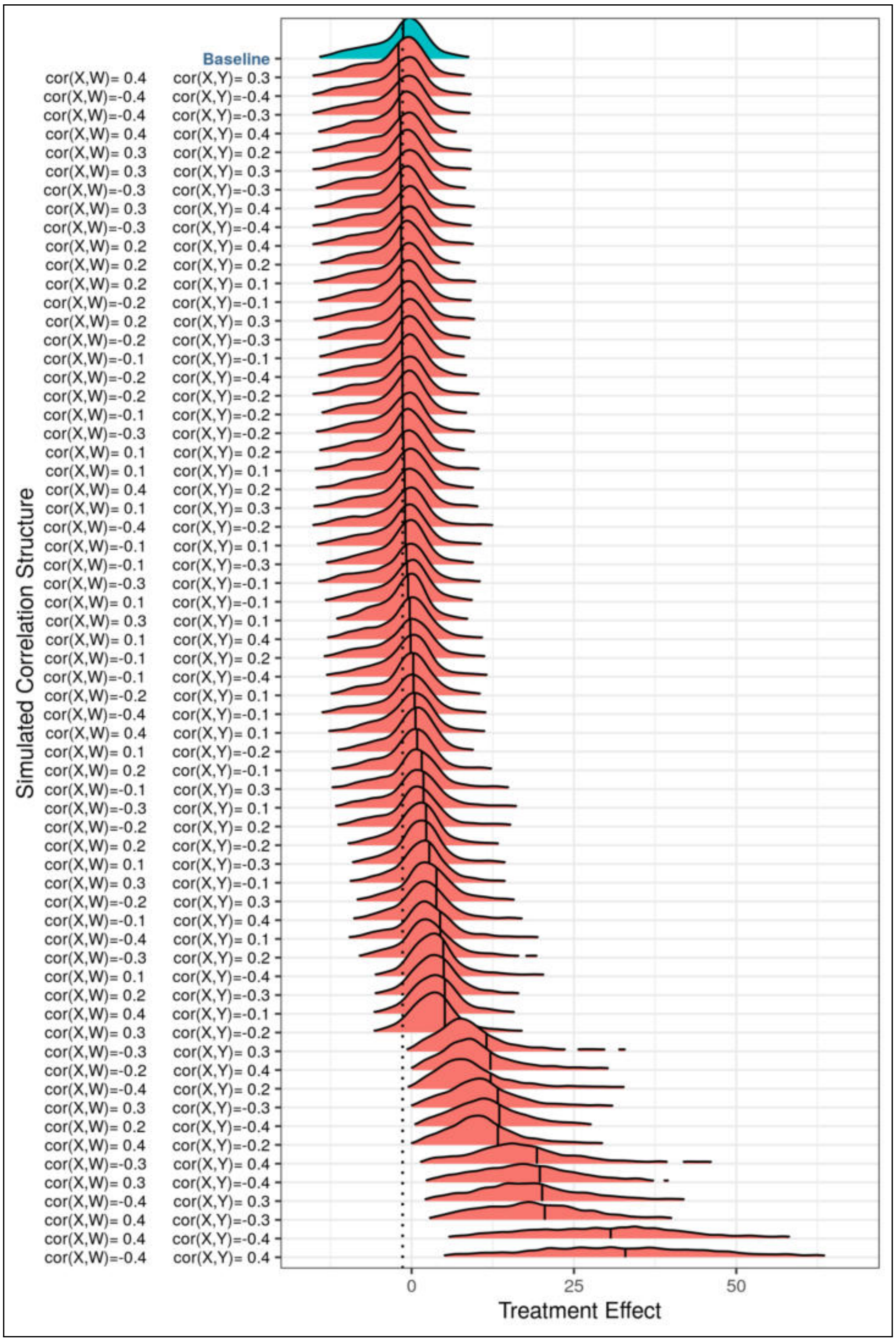


Figure 5.28: Pesticide intensity: Simulation of omitted variable with different correlation structures using the approach of Huber (2020). The corresponding treatment effect distributions are depicted and compared to the baseline model.

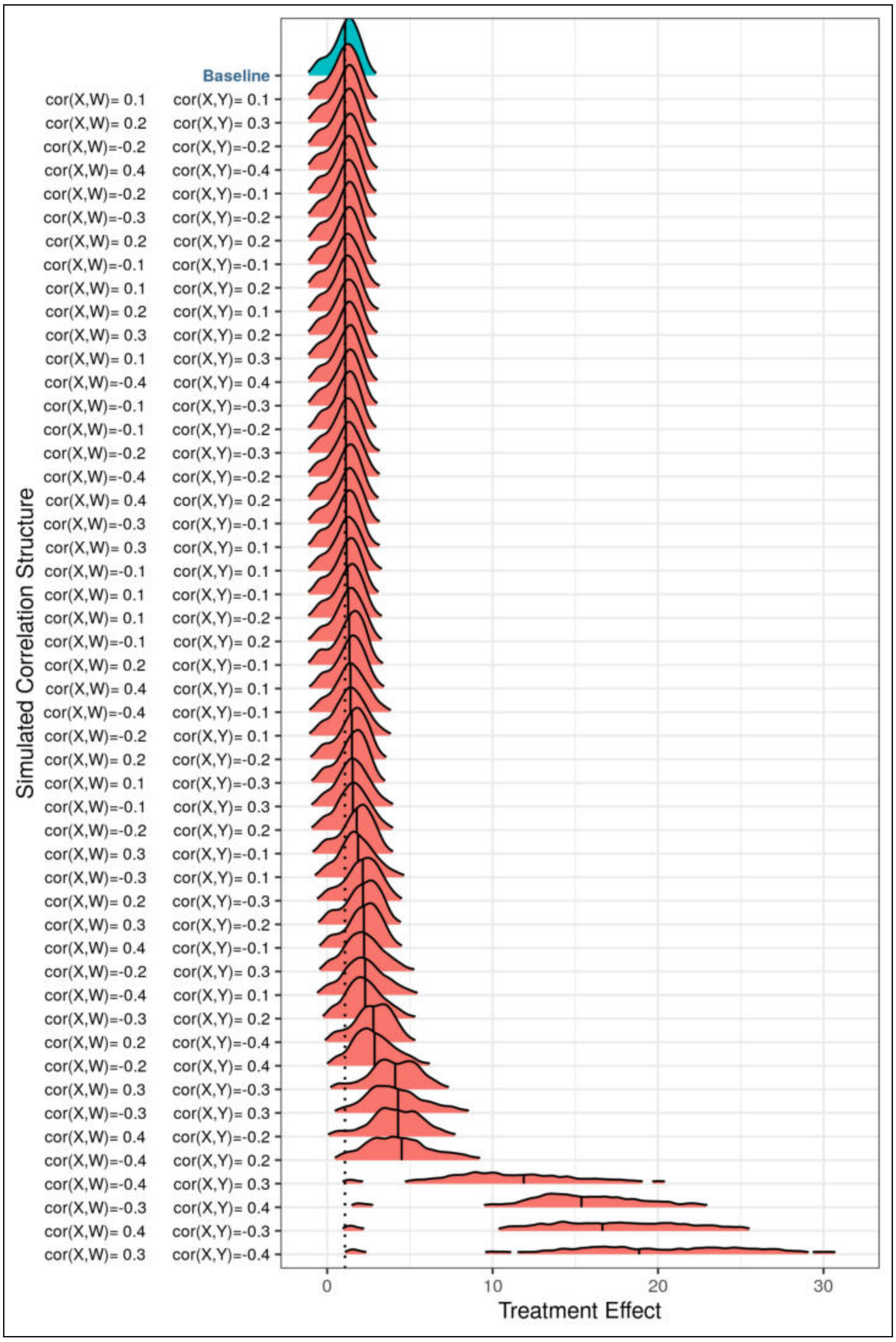


Figure 5.29: Land use diversity: Simulation of omitted variable with different correlation structures using the approach of Huber (2020). The corresponding treatment effect distributions are depicted and compared to the baseline model.

5.8.16 Robustness check: Observations outside baseline 95% confidence interval

Table 5.6: This figure shows the percentage of observations in the robustness checks that lie outside the 95 per cent confidence interval of the baseline model.

(Percentage of obs. outside 95% CI of baseline)	Greenhouse gas emissions	Fertilizer intensity	Pesticide intensity	Land use diversity
Leave out most important feature	0	0	0	0
Leave out 3 most important features	0	0	0	0
Random common cause	0	0	0.1	0.1
Leave out PC 1	0	0	0.2	11.6
Leave out PC 1-3	5.2	3.5	7.3	56.3
Leave out PC 1-6	10	9.1	14.9	48.9

6 Study IV – Tackling climate change: Agroforestry adoption in the face of regional weather extremes

Disclaimer:

This manuscript is coauthored with Johannes Sauer and currently under review at Environmental & Resource Economics. This paper was presented at the 96th Annual Conference of the Agricultural Economics Society (March 2022) in Leuven, Belgium.

Acknowledgments

The authors thank the members of the BLIZ research project consortium for their thoughtful comments. This study was funded by the Bavarian Ministry of Science and the Arts in the context of Bavarian Climate Research Network (bayklif).

6.1 Abstract

The cultivation of agroforestry systems is regarded as an effective strategy to synergistically mitigate and adapt to the impacts of climate change in the face of increased occurrences of regional extreme weather events. This study addresses the question of whether, and under what conditions, farmers are likely to adopt agroforestry and wood-based land-use systems in response to regional weather extremes and presents a novel research approach to tackle this question. A discrete choice experiment was conducted to elicit farmers' preferences for, and willingness to adopt, agroforestry and wood-based land use systems. The results were then combined with geospatial weather information. Assuming adaptive weather expectations, land users' dynamic responses to extreme weather years were simulated in terms of adoption probabilities. Farmers in the case study region in southeastern Germany were found to have a negative preference for alley-cropping systems (i.e. agroforestry) and short rotation coppice, compared to an exclusively crop-based land use system. However, the results from the simulation of a 2018-like extreme weather year showed that alley-cropping might have a very high probability of being adopted in the medium to long-term under different circumstances, thus enhancing farmers' resilience to climate change.

6.2 Introduction

The latest assessment report from the Intergovernmental Panel on Climate Change (IPCC) reiterates the fact that climate change poses exceptional challenges to various social and economic sectors on a global scale (IPCC, 2021). The World Economic Forum, in its 2020 Global Risks Report, listed climate-related concerns among the top-five long-term risks for the first time (WEF, 2020). In addition to affecting annual mean temperatures and precipitation, climate change also increases the number of occurrences of regional extreme weather events such as drought, heat waves, heavy rain and floods (IPCC, 2021; Lüttger & Feike, 2018; Mann et al., 2018; Westra et al., 2014). In this context, agriculture is often seen as one of the most susceptible sectors to such changes (IPCC, 2007), which negatively affect, e.g., crop yields (e.g. Haqiqi et al., 2021; Lesk et al., 2016; Schlenker & Roberts, 2009), total factor productivity (e.g. Chambers & Pieralli, 2020; Chambers et al., 2020; Stetter & Sauer, 2021), and ultimately farm income and viability (e.g. Dalhaus et al., 2020; Dell et al., 2014; Kawasaki & Uchida, 2016). Prime examples of years with extreme weather conditions, which heavily impacted agriculture are the 2003 European heat wave, the 2018 European drought and heat wave, and the 2010–2013 Southern United States and Mexico drought. Despite these impacts, agriculture is also regarded as one of the most important anthropogenic contributors to climate change (Lynch et al., 2021). Overall, farmers require effective adaptation and mitigation strategies to effectively deal with the challenges of climate change.

One major channel through which agriculture can actively tackle climate change impacts is land use (Pielke, 2005). A promising pathway in this direction is the adoption of agroforestry and wood-based land-use systems, which are recognized to play a key role in synergistically approaching adaptation and mitigation (Cardinael et al., 2021; Duguma et al., 2014; van Noordwijk et al., 2014, 2011; Verchot et al., 2007). Agroforestry systems are defined as land-use systems in which woody perennials are deliberately integrated with agricultural crops and/or livestock on a piece of land, either in some sort of spatial arrangement or in temporal sequence (Cardinael et al., 2021; Leakey, 2017; Nair, 1985). These systems mitigate climate change impacts through their carbon sequestration potential aboveground, belowground, and in the soil (e.g. Albrecht & Kandji, 2003; Cardinael et al., 2017; Oelbermann et al., 2004; Schroeder, 1993). There are also indirect mitigation effects created by the planting of trees and other woody perennials on agricultural land and these may also effectively reduce deforestation (Schroeder, 1993; Verchot et al., 2007) and help replace fossil fuels with fuel wood (Kuersten & Burschel, 1993). Simultaneously, the resulting positive regulation effects on hydrological cycles, soil, and the microclimate, may lead to more climate change resilient agricultural production practices (Lasco et al., 2014). Furthermore, agroforestry and its provision of multiple ecosystem services (Brown et al., 2018; Wolz et al., 2018) is also seen as a main component in the realm of ecosystem-based climate change adaptation (Hernández-Morcillo et al., 2018; Pramova et al., 2012).

Provided the various merits of agroforestry systems, there remains a large untapped potential for the introduction of agroforestry and its expansion across the globe (van Noordwijk et al., 2014), and this may become of even more importance in the face of an increased frequency of regional extreme weather events (Duguma et al., 2014; van Noordwijk et al., 2021).

This paper addresses the question of whether, and under what conditions, farmers are likely to adopt agroforestry and wood-based land-use systems in response to regional weather extremes. To answer this question, we conducted a discrete choice experiment to elicit farmers' preferences for, and willingness to adopt alley cropping (AC) and short rotation coppice (SRC). We combined the results with geo-spatial weather data. Assuming adaptive weather expectations, the farmers' dynamic land-use responses to extreme weather years in terms of adoption probabilities were locally simulated utilizing the approach of Ramsey et al. (2021). The paper then discusses these land-use responses in a wider climate change resilience context (see Meuwissen et al., 2019; OECD, 2020).

We found that farmers in our case study region of southeastern Germany have a negative preference and willingness to adopt (WTA) for AC and SRC compared to an exclusively crop-based land use system. However, the results from the simulation of extreme weather under different scenarios show that AC systems (agroforestry) might have a very high probability of being adopted in the medium to long-term and thus strengthening farmers' resilience to extreme weather events and climate change.

The elicitation of farmers' preferences for agroforestry and woody perennials has been the subject of multiple studies, including Gillich et al. (2019) and Pröbstl-Haider et al. (2016), who analyzed farmers' preferences for SRC in Germany and Austria using discrete choice experiments. Other studies focused on the adoption of agroforestry systems, mostly in the context of the Global South (Amusa & Simonyan, 2018; Bayard et al., 2007; Beyene et al., 2019; Dhakal et al., 2015; McGinty et al., 2008; Schaafsma et al., 2019). However, none of these studies, examined the effects of climate and (extreme) weather in this context. Furthermore, multiple authors have simulated the (economic) potential of agroforestry cultivation under different circumstances (e.g Frey et al., 2013; Paul et al., 2017). Lasch et al. (2010) and Gomes et al. (2020), including projecting the cultivation potential for SRC in eastern Germany and coffee-agroforestry in Brazil, all taking into account various climate change scenarios until 2050. The problem with such scenarios is that they are usually conducted on a global scale and likely do not represent local farmers' actual and perceived experiences with extreme weather and climate change. This is why they are usually not well-suited for farm-level based simulations (Morton et al., 2015; Ramsey et al., 2021). Overall, studies on the effects of weather shocks on land-use change are scarce (Girard et al., 2021). To fill in this gap, Ramsey et al. (2021) developed a novel framework to simulate how farmers dynamically adjust their cropping decisions in response to specific weather patterns.

This study contributes to the literature in several ways. It quantifies the link between adverse weather and farmers' preferences for agroforestry and SRC accounting for short- to long-term adaptation responses. While many of the aforementioned studies are concerned with why integrating woody perennials into farms' cultivation plans might be useful in terms of mitigation and adaptation, they frequently ignore whether and how farmers respond to weather patterns. Establishing this link is particularly important in light of the increased frequency of extreme weather events resulting from climate change. Furthermore, by combining a discrete choice experiment, geo-spatial weather information, and the simulation framework of Ramsey et al. (2021), this study is able to provide novel insights into farmers' responses and resilience in the face of a changing climate. By extending the work of Ramsey et al. (2021), the approach presented in this paper allows to evaluate two important aspects, that have widely been neglected in the literature on the climate change land use nexus so far. First, this nexus has usually only been studied *ex-post* for already established land uses. With the approach of this study, it is possible to evaluate this relationship for more recent, not yet established land use systems. Second, by taking choice-specific attributes into account, it is possible to simulate multiple alternative scenarios reflecting the role of legislation, market conditions, and technological progress in the context of climate change adaptation. Finally, the empirical case study focused on southeast Germany sheds more light on the adoption potential of agroforestry in the context of an industrialized country, since much of the work on this topic has until now been done in the context of developing countries.

The remainder of the article is structured as follows. First, a short description of agroforestry and wood-based agricultural land use systems is provided before presenting a conceptual framework (Sec. 6.3). In Section 6.4, the data collection and empirical strategy is presented. Section 6.5 describes the result from the discrete choice experiment and the weather simulations, followed by a discussion of the most important finding (Sec. 6.6). The paper closes with a summary and several concluding remarks in Sec. 6.7.

6.3 Background and conceptual framework

6.3.1 A short description of agroforestry and wood-based agricultural land use systems

As mentioned previously, agroforestry systems are land-use systems where woody perennials are integrated with agricultural crops and/or livestock on a piece of land, either in a spatial arrangement or in a temporal sequence (Cardinael et al., 2021; Leakey, 2017; Nair, 1985). This definition includes a wide range of diverse systems including silvopastoral (the combination of trees with livestock), silvoarable (planting crops between rows of trees), forest farming (food, herbal, botanical, or decorative crops under a forest canopy), home gardens; as well as hedge, windbreak, and riparian buffer strip systems and many more (Pantera et al., 2021; USDA, 2019). As can be seen by this diverse list, agroforestry is not a new concept and goes back a very long time in many regions of the world (Pantera et al., 2021).

With regard to the integration of trees on agricultural land, SRC have been identified as an attractive land-use alternative from both economic and ecological perspectives (Baum et al., 2009; Wolbert-Haverkamp & Musshoff, 2014). SRCs usually consist of fast-growing tree species such as poplar, willow, paulownia, robinia, or eucalyptus with short rotation periods and frequent harvests (every 3 to 5 years) (Rödl, 2017). Other than agroforestry, SRCs are typically associated with a single use on the same field.

More recently, AC systems that integrate strips of SRCs into agricultural fields have been receiving increasing attention (Tsonkova et al., 2012). In this system, farmers produce crops and woody biomass on the same field at the same time. This can result in multiple advantages across several domains.

Several previous studies, including Paul et al. (2017), Gosling et al. (2020) and Schoeneberger et al. (2017) have found that AC can generate higher economic returns than single crop land uses. Furthermore, diversifying production output can raise economic stability (Tsonkova et al., 2012), and ACs can contribute to a more sustainable biobased economy by simultaneously providing food and renewable raw materials (Gillich et al., 2019). Numerous studies have also found positive effects on crop yield and land-use efficiency (see e.g. Schoeneberger et al., 2017).

AC also provides a range of environmental services due to its multi-functional nature. It has the ability to break up large-scale structures, increase biodiversity through increased habitat, increase species diversity and their connectivity throughout agricultural landscapes, and it can also reduce soil erosion and nutrient leaching (Langenberg & Theuvsen, 2018; Schoeneberger et al., 2017; Tsonkova et al., 2012).

Finally, agroforestry systems and SRC can, to some degree, play an important role in synergistically approaching climate change mitigation and adaptation. In terms of climate change mitigation, AC and SRC systems can store large amounts of carbon in aboveground and belowground biomass (Albrecht & Kandji, 2003) as well as in soil (Cardinael et al., 2017), thus reducing atmospheric carbon dioxide (Cardinael et al., 2021; Schroeder, 1993; Tsonkova et al., 2012). Regarding adaptation, the integration of trees on agricultural land provides a buffer to weather extremes through regulating hydrological cycles, improving nutrient and water-use efficiencies, and modifying microclimates (Ashraf et al., 2019; Pramova et al., 2012; Wolz et al., 2018). Agroforestry can also diversify farmer income by hedging financial risk (Wolz et al., 2018), and can make production more resilient to the negative effects of climate change (van Noordwijk et al., 2021).

Despite these myriad advantages, silvoarable agroforestry systems are still relatively rare in Europe (den Herder et al., 2015; Langenberg & Theuvsen, 2018). van Noordwijk et al. (2014) note that there is a huge potential for the introduction and expansion of agroforestry areas around the globe.

6.3.2 Land-use, random utility maximization and weather expectations

Given the large potential for the introduction and expansion of agroforestry, this study seeks to elicit farmers' preferences for agroforestry and SRC in comparison to conventional crop farming against a background of climate change. The theoretical basis for this analysis is based on random utility maximization following Lancaster (1966) and McFadden (1973). When it comes to planning the usage of their land, farmers face a choice among a set of alternative land uses in various decision situations under varying conditions. Each farmer obtains a certain level of indirect utility from each land-use alternative. In a given decision situation t, she will select alternative i if and only if $U_{it} > U_{jt}, j \neq i$. The indirect utility of an alternative cannot be directly measured but it can be expressed by a systematic (deterministic) component V, reflecting specific characteristics as well as farmers' individual and location-specific features, plus a random component ϵ, representing unobserved decision-relevant elements (Mariel et al., 2021). A farmer n obtains a certain level of indirect utility U_{njt} from a land use alternative j in a choice situation t.

$$U_{njt} = V_{njt} + \epsilon_{njt} \tag{6.1}$$

As is standard, it is assumed that farmers' utility for a land-use alternative to vary with a set of decision-relevant characteristics (x, see Sec. 6.4.2). Furthermore, as agricultural land-use is heavily dependent on weather (c), it is assumed that farmers' utility also depends on expected weather at the time of the planting decision:

$$V_{njt} = f(x_{njt}, c_{nt}; \beta, \gamma) \tag{6.2}$$

where β and γ are coefficients to be estimated. Following Nerlove (1958) and Ramsey et al. (2021), it is also assumed that farmers have adaptive weather expectations that are based on past local weather history, where both short-term and long-term trends might affect land use choices. What is more, it is realistic to assume that farmers do not assign equal importance on each past weather event, which is why a simple average of past weather would not properly reflect farmers' expectations. Ramsey et al. (2021) express the expected-weather-formation-process as follows:

$$c_{nkt} = \omega_0 + \omega_s\, W(\ddot{c}_{nkt-1}, \ldots, \ddot{c}_{nkt-r}) + \omega_l\, W(\ddot{c}_{nkt-r-1}, \ldots, \ddot{c}_{nkt-R}) \tag{6.3}$$

where $\ddot{c}_{nkt-r}$ are actual past weather events. ω_0 is a reference expectation, ω_s reflects a farmer's weight assigned to the recent past, ω_l is the weight assigned to the more distant past, and $W(\cdot)$ is a weighting function (e.g. annual mean). Hence, weather expectations are formed by two components, one reflecting longer term weather patterns ("signal") and one reflecting short term weather variations ("noise"). In terms of climate change adaptation (i.e. the adoption of novel land use options), one may presume that the signal plays the dominant role in decision making. However, especially with respect to severe, more tangible weather events, the noise component might be more important because of its immediate negative effect on production (Ramsey et al., 2021), while such an experience might level off with ordinary weather events in the longer-term.

In light of these theoretical considerations, past (extreme) weather events are expected to influence farmers' decisions to adopt more climate change resilient land-use options such as agroforestry or SRC .

6.4 Material and methods

This section first provides information on the case study region, Bavaria, before describing the discrete choice experiment (DCE) setup used to collect data on farmers' preferences. Then the data that are used to describe weather are presented. By combining the experimental with the weather data and utilizing a correlated random parameter logit (RPL) approach, it is possible to estimate farmers' preferences and probabilities for the cultivation of each land-use option and to retrieve coefficient estimates reflecting the influence of land-use characteristics and (anticipated) weather. Finally, the simulation approach used to model the adaptive adjustment behavior of farmers in response to an extreme weather year based on the estimates from the RPL model is described.

6.4.1 Study area

The DCE was conducted in Bavaria, a federal state of Germany in Central Europe. Located in the southeast of Germany, Bavaria belongs to the core regions of agricultural production within the European Union (EU). It reflects the variety of European farming (conditions) to a high extent, which is why this site was selected for conducting this study. In terms of natural conditions, farming takes place along an elevational gradient of 1500 m (from 100 m in Northwest Bavaria to 1600 m in Southeast Bavaria) and a macroclimatic gradient with a mean annual temperature range between 3 °C and 10 °C and an annual precipitation of 470–1592 mm (from 1960 to 2020). Its natural conditions range from pre-alpine and alpine areas with high precipitation and rather clayey, limestone and dolomite based soils to regions with flat land and fertile loess soils to dry, marlstone, limestone and dolomite based hillside locations. They are well-suited for various agricultural production systems including crop farming, intensive and extensive dairy farming, pig and cattle fattening and breeding, poultry farming, vegetable farming, orcharding, hop production and viticulture. This heterogeneity is to a high degree reflected in Bavaria's seven regional districts (Figure 6.1). These will be analyzed individually in addition to the entire region in the results section. Appendix 6.8.1 provides a detailed description of the structural and natural conditions for each district.

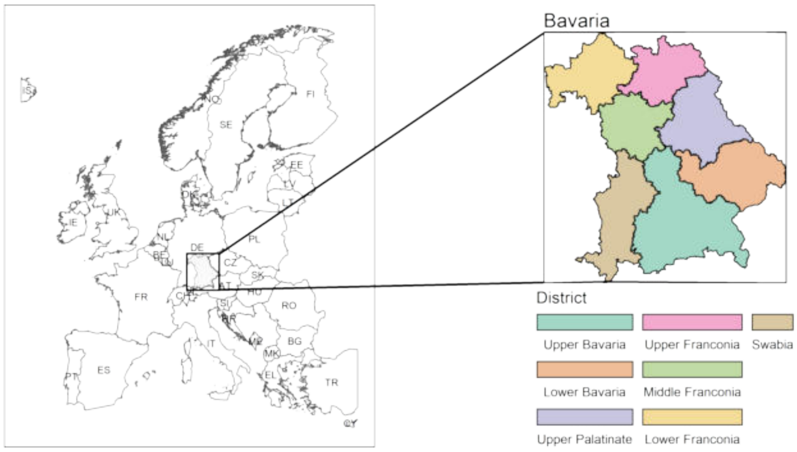


Figure 6.1: The case study region Bavaria is a federal state of Germany and lies in Central Europe. It is comprised of seven regional districts.

6.4.2 Choice experiment setup

A DCE was used to elicit the influence of land use characteristics on farmers' decisions regarding whether or not to cultivate agroforestry. Each farmer was repeatedly confronted with a choice situation, in which the attributes of three land-use alternatives (namely SRC, AC, and status quo crop farming) varied.

Attribute selection and levels

Following a careful literature search and after receiving feedback from agricultural experts, the attributes used to describe the land-use alternatives are the following: average yearly margin contribution, yearly margin contribution variability, minimum useful lifetime, payments for ecosystem services and a dummy if the alternative qualifies as ecological priority area.

The primary monetary attribute, the margin contribution, measures yearly revenues (yield times price), minus variable cost. Fixed cost and subsidies are not considered in this measure. Moreover, because revenues and costs are spread over the entire production period of SRC and AC, a margin contribution equivalent, which corresponds to the annualized form of the net present value is introduced.

Previous studies show that uncertainty plays an important role when it comes to farmers' decision making processes in general (see e.g. Menapace et al., 2013) and land allocation in particular (El-Nazer & McCarl, 1986; Knoke et al., 2015). This study expresses outcome uncertainty in terms of gross margin fluctuations. Since farmers tend to be risk-averse (Menapace et al., 2013), we expect an increase in variability to negatively affect preferences.

The minimum useful lifetime of a land-use alternative is closely related to the entrepreneurial flexibility of farm businesses. Being tied longer-term to one land-use type means a loss of flexibility (Musshoff, 2012). This is expected to negatively affect farmers' preferences.

Since SRC and AC provide a wide range of environmental services, payments for ecosystem services (PES) could provide a positive incentive for farmers to cultivate one of these land-use options (e.g. Layton & Siikamäki, 2009).

Finally, Langenberg et al. (2018) find that one major driver for farmers to engage in AC may be the area's qualification as "ecological priority area". To achieve this designation, farmers must attribute a certain amount of land to ecological priority areas (which are considered environmentally-friendly). They then receive area based "greening" payments, which account for approximately 30% of the farmers' total basic payment.

We aimed for realistic levels of each attribute based on official databases (e.g. LfL, 2018; StMELF, 2018), previous studies (e.g. Gillich et al., 2019; Hauk et al., 2014; Langenberg et al., 2018; Pröbstl-Haider et al., 2016), expert consultation and plausibility considerations. Table 6.1 summarizes the attributes and attribute levels.

Table 6.1: Description of attributes and levels.

Attribute	Description	Attribute levels
MC	Margin contribution (equivalent) (€/ha)	400[a], 600, 800
MCV	Margin contribution variation (%)	15[a], 30
MUL	Minimum useful lifetime (years)	3[a,b], 16, 20, 24
PES	Payment for environmental services (€/ha)	0[a], 100, 200
Green	Qualification as ecological priority area	Yes, No[a]

[a] Fixed attribute levels for the status quo alternative.
[b] Attribute level that only applies for the status quo alternative.

Conducting the choice experiment

After having determined choices, attributes and corresponding levels the actual choice experiment commenced. A choice experiment with three labeled alternatives was created, namely "Short Rotation Coppice", "Agroforestry", and "Status Quo". Given its labeled nature, we followed Viney et al. (2005) and created an L^{MA} design for the DCE, that resulted in 36 choice cards. To reduce the psychological burden of answering all choice tasks, they were randomly blocked into three sets of twelve choice cards. An exemplary choice card can be found in appendix 6.8.3. Each respondent was then randomly assigned to one of the three blocks. Before the participants started the choice experiment, an explanation of how the DCE would work was provided, along with descriptions of the alternatives and attributes relevant for the task (see appendix 6.8.4). To tackle hypothetical bias, we used "cheap talk" (Landry & List, 2007) and reminded the participants about the danger of hypothetical bias and that they should answer truthfully.

The survey consisted of several parts. After some general information and the respondents' consent to participate, they were asked for general (socioeconomic) characteristics of their farm, which was followed by the DCE. Finally, the participants were asked to give further information on their local climate change perception and several character traits.

After an extensive pre-test phase which took place in the early summer, the survey was conducted online in October 2020. Respondents from Bavaria were recruited from a large panel of farmers provided by an agricultural market research platform called agriEXPERTS and through multiple outlets of a specialist publishing house for agriculture (`Deutscher Landwirtschaftsverlag, dlv`). The survey included an invitation to take part in a lottery to win one of ten vouchers for a popular agricultural clothing shop worth 50 EUR each. It took approximately twelve minutes to complete the questionnaire.

6.4.3 Weather variables

To accurately describe the local weather history of these farms, five common weather indicators were selected, including average temperature, precipitation sum, number of dry days, number of hot days and the number of heavy rain days, during the local growing season (March–October). The variables are derived from 0.1 degree gridded daily data from the European Climate Assessment & Dataset (ECA&D) project (Cornes et al., 2018). Following ETCCDI (2018) and DWD (2022), dry days are defined as days with precipitation of < 1mm and hot days are defined as days with maximum temperature > 30 °C. On heavy rain days precipitation exceeds 20 mm (DWD, 2022). The weather indicators were aggre-

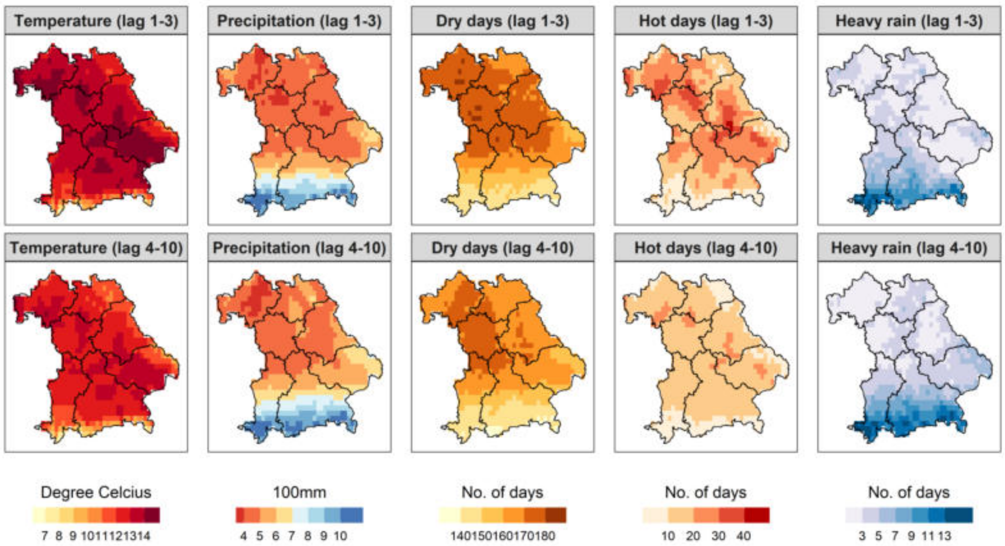


Figure 6.2: Summary of the weather variables used for the estimation of the baseline model with lag structure 1-3 and 4-10.

gated at the municipality level (2031 municipalities) and linked to the responses from the questionnaire via zip codes.

As outlined in Section 6.3.2, farmers form their weather expectations based on historical weather patterns, which can be distinguished between short-term and longer-term weather patterns. To capture this distinction, short-term and long-term weather variables were defined by different lag structures. The base specification for short-term weather patterns for the five indicators were based on the average of years $t-1$ to $t-3$ (more recent past) and longer-term weather patterns are based on the average of years $t-4$ to $t-10$ (more distant past). Figure 6.2 summarizes these variables. Further lag structures were computed reflecting multiple candidate time horizons of expectation formation, which were later tested against the base structure (sec. 6.5.2). All weather variables were mean-centered, which will prove useful for the interpretation of the alternative-specific constants in the RPL model and the corresponding willingness to adopt measures.[1]

[1] Mean-centering the weather variables allow to directly interpret these measures, as the intercepts are evaluated at the average weather (for which all weather variables take on a value of zero.)

6.4.4 Econometric approach

For the econometric analysis, the random parameter logit model was used to account for preference heterogeneity in the utility function of the investigated sample (Hensher et al., 2015; Train, 2001). The utility function (6.2) was parametrized with alternative specific constants (α_{ij}), land-use specific attributes (X) and individual-specific weather parameters (C):

$$V_{ijt} = \alpha_{ij} + \beta_i X_{jt} + \delta_{ij} C_i + \epsilon_{ijt} \tag{6.4}$$

The model formulation is a one level multinomial logit model, for individuals $i = 1, \ldots, N$ in choice setting t and alternatives j. It assumes a Gumbel distribution of the error term ϵ_{ijt}, and the probability of each choice j is as follows (Hensher et al., 2015):

$$Prob(y_{it} = j) = \frac{\exp \alpha_{ij} + \beta_i X_{jt} + \delta_{ij} C_i}{\sum_j (\exp \alpha_{ij} + \beta_i X_{jt} + \delta_{ij} C_i)} \tag{6.5}$$

In the RPL framework, the coefficient vectors α_{ij}, β_i, and δ_{ij} are considered random draws from a distribution whose parameters need to be estimated. Under this assumption, we can use maximum simulated likelihood estimation to obtain coefficient estimates for α_{ij}, β_i, and δ_{ij} (Train, 2001). A total of 1000 Halton draws were used for each model estimation. As for the parameter distributions, it was assumed:

$$(\alpha_{ij}, \beta_i) = (\alpha_j, \beta) + \Gamma \; \nu_{(j)i} \tag{6.6}$$

$$\delta_{ij} = \delta_{ij} + \Omega \; \upsilon_{ij} \tag{6.7}$$

where $\nu_{i(j)}$ and υ_{ij} describe random unobserved preference variation, with mean zero and covariance matrix with known values on the diagonal, and fixed by identification restrictions. Γ is a lower triangular matrix that allows correlation across the attribute-related random parameters and $\Omega = diag(\sigma_1, \ldots, \sigma_k)$.

This specification followed the approaches of Hess & Rose (2012) and Hess & Train (2017), who showed that only by allowing for correlation across attribute-related random parameters, is it possible to capture scale heterogeneity alongside heterogeneity in utility coefficients. According to Hess & Rose (2012, p.9): "Such correlations can be expected in any setting: they simply reflect that respondents' preferences for one attribute are related to their preferences for another attribute". Ignoring this correlation could severely bias parameter estimates. One reason why this specification is only rarely observed in the literature might be its significantly higher computational burden (Mariel & Meyerhoff, 2018).

It is assumed that all random parameters were normally distributed except for the coefficient of the contribution margin, which was assumed to be log-normally distributed. This was done for two reasons. First, economic theory states that the sign for the profit attribute should always be positive. Second, finite moments for the WTA values (Daly et al., 2012) are assured. These are defined as the change in one attribute with respect to the return margin. Hence they are the ratio of each parameter estimate and the parameter estimate of the marginal contribution:

$$WTA = \frac{(\widehat{\alpha}_{ij}, \widehat{\beta}_i)}{\widehat{\beta}_{i,\text{contribution margin}}} \tag{6.8}$$

By mean-centering the weather variables we can make direct use of the estimated ASCs. Dividing them by the individual-specific coefficient estimate on the return margin gives the marginal WTA at mean weather because all mean centered weather variables in C are zero at their means (see also Iacobucci et al., 2016).

6.4.5 Post-estimation simulation

To evaluate the short- and longer-term adjustment dynamics to an extreme weather period, farm-level responses to one to five-year weather shocks over a period of 10 years were simulated following the method of Ramsey et al. (2021). The simulation is based on the estimated parameters from the fitted RPL model in Sec. 6.4.4. These simulations are primarily based on the 2018 drought year, which caused severe damage to German crop farming (Webber et al., 2020). Following the reasoning of Girard et al. (2021), that different weather shocks have different impacts on land-use responses, we also present simulation results for a 2003-like heatwave (Ciais et al., 2005).

Given the lag structure of the weather variables, farmers' adoption probabilities in response to an extreme weather period each year during and after the weather shock based on the formula for land use probabilities can be simulated (Eq. 6.5). In the baseline scenario, the values of the weather variables C were replaced for every farm in years 0–10 with their respective (sub-)sample long-term averages (LTA) over the 20-year period 1991–2020. For a one-year shock scenario, the 2018 (2003)-like event is assumed to occur in period $t = 0$, and then the weather returns to the LTA. This shock will affect the values for the short-term weather variables (lags 1–3) in periods 1–3, and then they return to the LTAs. The longer-term weather variables (lags 4–10) remain at the LTAs for periods 1–3 before changing to a "shocked" level in years 4–10 after the shock (compare Ramsey et al., 2021, p.13 and App. 6.8.5). Figure 6.3 illustrates the composition of each weather variable over time for a one-year, a two-year and a three-year weather shock as they enter equation 6.5 in the simulation.

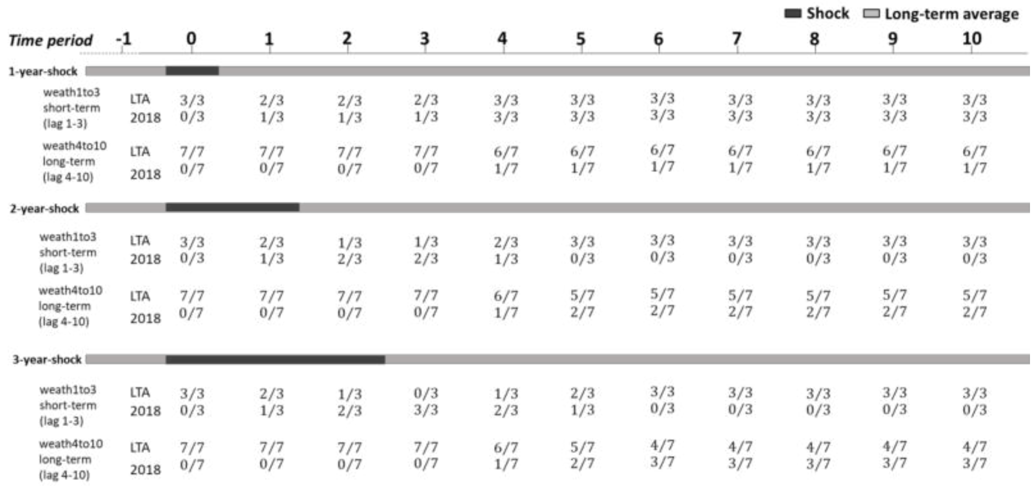


Figure 6.3: Illustration of the composition of the weather variables as they enter the simulation scenarios and replace the original weather variables used for the RPL estimation. The replacement procedure is demonstrated of a one-year, two-year and three-year shock scenario. Longer-term shocks change accordingly.

We ran simulations for the full sample as well as for each district separately to explore more potential regional adaptation paths. Table 6.2 summarizes the respective values used for the construction of the weather variables.

Regarding the levels of the land-use attributes X, we constructed several scenarios, reflected by different attribute levels used in the simulations. The respective levels and scenarios are summarized in Table 6.3.

6.5 Results

6.5.1 Sample summary statistics

In total, we received 210 responses. Twelve responses were deleted after a series of plausibility checks. The resulting analysis is thus based on the responses of 198 farmers. In Table 6.4, summary statistics for key farm characteristics in this sample are described and compared with the population means for Bavaria (mostly stemming from official census data). Approximately half of the sample is full-time

Table 6.2: Description of the weather indicators as they enter the 2018-like shock simulations.

	Precipitation (mm/year)	Average Temp. (°C)	Dry days	Heavy rain days	Hot days
Bavaria (full sample)					
Long-term average (baseline)	610.21	12.39	156.35	4.92	12.36
Extreme weather year (2018)	421.52	13.93	181.14	3.75	24.63
Difference	-188.70	1.53	24.79	-1.17	12.26
Upper Bavaria					
Long-term average (baseline)	778.48	12.27	147.98	7.77	11.14
Extreme weather year (2018)	583.21	13.70	170.29	6.23	17.99
Difference	-195.27	1.44	22.32	-1.54	6.85
Lower Bavaria					
Long-term average (baseline)	601.28	12.57	156.37	4.48	13.09
Extreme weather year (2018)	421.10	14.27	180.66	3.34	26.73
Difference	-180.18	1.70	24.29	-1.14	13.64
Upper Palatinate					
Long-term average (baseline)	509.76	12.30	160.66	3.09	12.52
Extreme weather year (2018)	327.00	13.97	184.70	1.97	29.04
Difference	-182.75	1.67	24.04	-1.12	16.52
Upper Franconia					
Long-term average (baseline)	525.48	12.13	158.80	3.40	11.58
Extreme weather year (2018)	302.93	13.72	185.51	1.60	25.63
Difference	-222.55	1.59	26.70	-1.80	14.05
Middle Franconia					
Long-term average (baseline)	474.35	12.85	164.70	2.91	15.00
Extreme weather year (2018)	316.06	14.39	193.00	2.93	31.40
Difference	-158.29	1.54	28.30	0.02	16.40
Lower Franconia					
Long-term average (baseline)	463.42	12.90	163.59	2.54	15.41
Extreme weather year (2018)	304.78	14.57	192.77	2.54	35.61
Difference	-158.64	1.67	29.18	0.00	20.20
Swabia					
Long-term average (baseline)	721.56	11.93	152.09	6.98	9.54
Extreme weather year (2018)	506.23	13.17	174.49	4.94	14.21
Difference	-215.33	1.24	22.41	-2.03	4.68

Table 6.3: Simulation scenarios and corresponding attribute values.

	Alley-Cropping					Short Rotation Coppice					Status Quo				
Scenario	MC	MCV	MUL	PES	Green	MC	MCV	MUL	PES	Green	MC	MCV	MUL	PES	Green
1. Regular-case	400	30.0	24	0	No	400	30.0	24	0	No	400	15	3	0	No
2. Regular-case w/ policy support	400	30.0	24	200	Yes	400	30.0	24	200	Yes	400	15	3	0	No
3. Regular-case w/ technological improvement	400	30.0	16	0	No	400	30.0	16	0	No	400	15	3	0	No
4. Better-case for agroforestry	600	22.5	20	100	Yes	600	22.5	20	100	Yes	400	15	3	0	No
5.Ideal-case for agroforestry	800	15.0	16	200	Yes	800	15.0	16	200	Yes	400	15	3	0	No

Note: MC = Margin contribution (Euro), MCV = margin contribution variation (%), MUL = Minimum useful lifetime (years), PES = Payments for environmental services (Euro), Green = Cultivated area eligible for greening premium

Table 6.4: Sample description and comparison with the population mean.

	Sample			Bavaria
	Mean	Median	SD	Population Mean
Full-time farming (1 if yes, 0 otherwise)	0.49	0	0.5	0.45[c]
Utilized area (ha)	69.79	40.5	89.84	36.66[c]
Share of cropland (%)	59.57	64.55	27.13	65.18[c]
Share of grassland (%)	34.07	30	22.9	34.37[c]
Share of forested land (%)	10.19	5	12.37	–
Share of rented land (%)	31.13	20	30.71	51.0[c]
Workforce (AWU[a])	0.1	0	0.3	0.12[c]
Full-time farming (1 if yes, 0 otherwise)	1.63	1.25	1.29	2.27[c]
Farmer's age (years)	48.34	50	12.42	50.3[d]
Higher education (1 if yes, 0 otherwise)[b]	0.24	0	0.43	–
Participation in agri-environmental program (1 if yes, 0 otherwise)	0.73	1	0.44	0.68[e]

Note: Number of observations = 198; a AWU denotes annual working units. b Higher education refers to having a university degree.

Sources: c Destatis (2021b), d LfL (2015), e Destatis (2021a)

farmers, which is only slightly higher than the Bavarian average (45%). Several characteristics of this sample are similar, on average, to the Bavarian average, namely cropland and grassland shares, farmers' ages and the participation rate in agri-environmental programs. At the same time, these sample farms manage more land on average, have a smaller share of rented land, and have a smaller workforce than the population mean. Also, the sample share of organic farms with 10% is very similar to the population share (12%). Overall, the descriptive statistics show that our sample reflects the Bavarian farmer population reasonably well, except for a few dimensions including farm size and labor. These deviations from the population mean are not necessarily negative in light of a dynamic trend toward fewer but larger farms within the EU (Wimmer & Sauer, 2020). Nearly all farmers stated they had already experienced negative consequences due to climate change related extreme weather events, especially in the form of yield and quality losses.

6.5.2 Model estimates and willingness to adopt

Model estimates

The model estimation results are summarized in Table 6.7. In a first step, we compared the model of our choice – the correlated RPL model – to a multinomial logit model (Model 1), and an uncorrelated RPL model with weather variables (Model 2). Likelihood-ratio tests showed that the correlated RPL model was a significantly better fit to the data than the alternative models. This is why this model was the preferred choice. Table 6.8 shows the parameters' correlation structure. Additionally, from these tests it was possible to empirically confirm that the weather (history) variables jointly have a significant impact on farmers' land use decisions as assumed in theory section. From Table 6.7 (Model 3), we can see by the attribute specific constants (ASC) that crop farming is preferred to both SRC and AC during average weather conditions. Preference heterogeneity (indicated by the estimates of the standard deviation of the random parameters) was also found for most choice attributes, except for margin contribution variability.

Willingness to adopt

To obtain further insights into farmers' land-use preferences, we calculated farmers' WTA based on the individual coefficient estimates from the correlated RPL. Figure 6.4 presents the results separately for both the full sample as well as for regional districts. Panel A shows that farmers had a generally negative WTA with respect to agroforestry (median value of −€123 for Bavaria) and SRC (median value of −€513 for Bavaria) evaluated at average weather conditions. These values can be interpreted as the farmers' *ceteris paribus* compensations to cultivate the corresponding alternative in addition to the contribution margin from the status quo crop rotation. SRC is valued more negatively and has a larger heterogeneity than AC (and crop farming), a pattern that is seen consistently across regional districts. Median WTA values for AC range from −€83 in Lower Bavaria to −€132 in Central Franconia, and for SRC from −€286 in Lower Bavaria to −€544 in Central Franconia. Hence, on average, farms in Central Franconia seem to be most adherent to status quo crop farming, while farms in Lower Bavaria seem to be most willing to switch to agroforestry and SRC.

When examining the land-use characteristics (Figure 6.4 panel B), it can be seen that an increase in the contribution margin variability (i.e. higher economic risk), as well as an increase in the minimum useful lifetime (i.e. lower entrepreneurial flexibility), decreases the willingness to adopt the land use options. This was the case for all regions to varying extents. This also meant that a reduction in these variables could lead to a higher willingness to cultivate agroforestry and SRC. For instance, the negative WTA for agroforestry in the Lower Franconia subsample (median: −€128) could potentially be offset by a *ceteris paribus* reduction of the minimum useful lifetime (median WTA: −€33.4) by 4 years. Furthermore, offering PES lead to an increase in the WTA. For example, within the full sample, every extra PES Euro leads to an €0.53 increase. While this value varies across observations and regions, the increase in ecosystem payments is in many cases under-proportional to the increases in the WTA (< 1) and thus rather inefficient. As already evident from the data presented in Table 6.7, designating land as an ecological priority area is not a useful lever to increase farmers' willingness to cultivate AC and SRC.

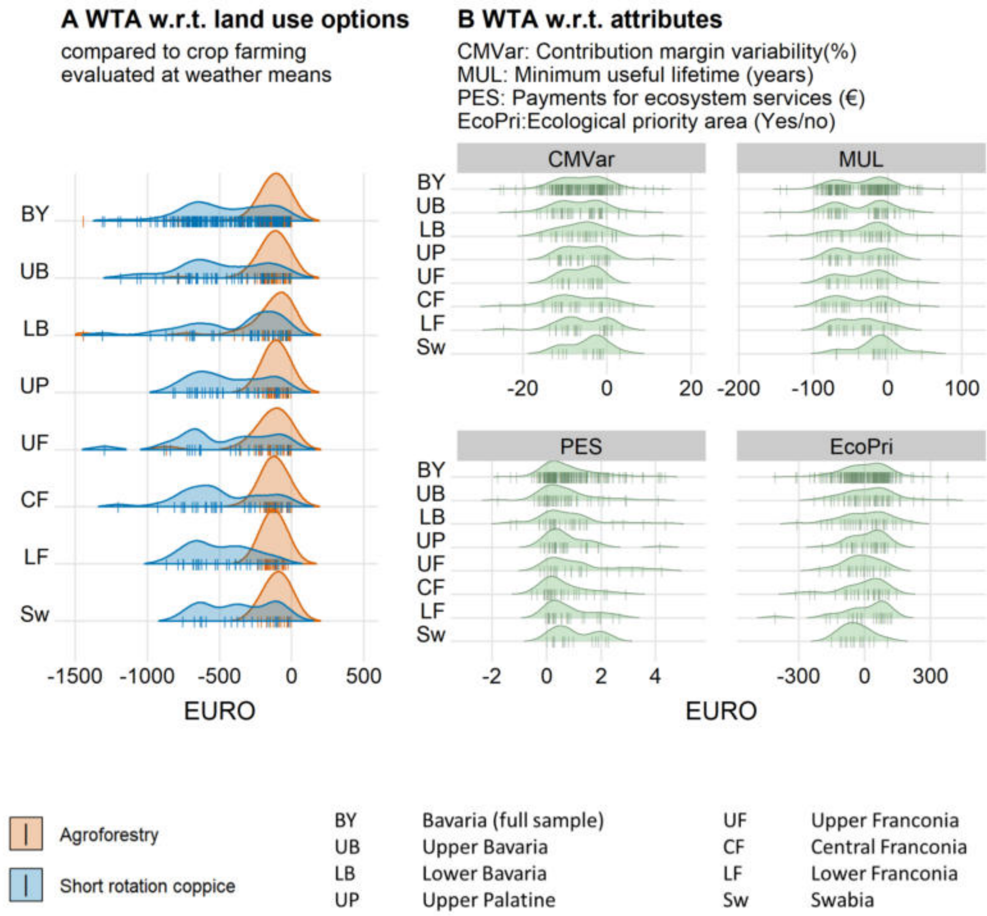


Figure 6.4: Summary of individual-specific willingness to adopt (WTA) estimates expressed as EUR/ha.

This paper refrains from analyzing the coefficient estimates of the weather variables individually because they likely suffer from a certain degree of multi-collinearity, which is not a problem *per se* but makes *ceteris paribus* statements very difficult. This issue is further examined in Sec. 6.6.

Further robustness checks

As mentioned previously, there are multiple possibilities for the empirical specification of the lag structure of the weather history data reflecting longer-term weather patterns ("signal") and short-term weather variations ("noise"). Therefore, we tested a series of alternative weather variable specifications and re-estimated the correlated RPL model and compared the model fit with our selected model (lags 1-3 and lags 4-10) (Table 6.9). It can be seen that the preferred model fits the data best followed by models with lag structures 1/2-15 and 1-3/4-15.

6.5.3 Weather simulations

To examine farmers' agroforestry adoption in response to the more extreme and adverse weather patterns, which are predicted to occur more often and last longer, we simulate a 2018-like (and 2003-like) extreme weather year at the regional level and observe the deviations of land-use probabilities from the average thirty-year baseline weather considering multiple socioeconomic scenarios. Additional simulations of the same weather events lasting for three and five years, respectively, were conducted. Furthermore, we developed an interactive simulation tool that allows to flexibly adjust and combine the simulation settings according to one's individual needs. This tool is available at: `https://ge36raw.shinyapps.io/main_dashboard/`.

Examining the pre-defined scenarios in Figure 6.5, for a one-year 2018-like weather shock, one can see that in the regular-case scenario (all land-use options at their base levels), farms' land-use probabilities remain close to their baseline levels after an adjustment period. In all regions except Swabia crop farming was the preferred land-use type. However, if the settings are changed such that AC and SRC experience policy support ("Regular-case w/ policy supp"), then AC becomes at least equally likely to be adopted as crop farming in Upper and Lower Bavaria. This effect is even more pronounced if the minimum useful lifetime of wood-based cultivars is reduced to sixteen years. From the scenarios with more preferential conditions for the wood-based land uses, it can be seen that farmers in Upper

Palatinate and Middle Franconia are quite reluctant to adopt these land use types. An interesting pattern can be observed across all scenarios and regions. In the first years after a one-year shock, farmers tend to prefer status quo crop farming, indicated by an increasing adoption probability. Similar outcomes for a 2003-like weather event (Figure 6.8) are also found, although less adjustment movement can be observed in the case of crop farming but more pronounced adjustment movement in the case of SRC. Also, we find that AC also becomes more likely to be adopted in Swabia.

To consider a longer duration shock, Figure 6.6 shows the simulation results for a three-year 2018-like extreme event. Over all, we find very similar patterns to those seen before. Nevertheless, following the more pronounced extreme weather event (in terms of duration) the farms' adaptation path following such an event becomes also more marked. For instance in Lower Bavaria, the probability of cultivating crops first reaches near 100% in scenarios one and three following the extended weather shock before falling considerably below the baseline level (and the probability of adopting AC). It was also found that without policy intervention or the shortening of the minimum useful lifetime, AC becomes the preferred land-use option in both Upper and Lower Bavaria. In a scenario of a 2003-like extended weather event (Figure 6.9) and when examining the full sample, AC becomes the preferred land-use choice regardless of the scenario. Additionally, the adaptation path regarding crop farming becomes considerably more volatile.

Finally, examining an even more extensive weather shock sceanrio, which spans five years (Figure 6.7), it can be seen that AC eventually becomes the preferred land-use type in almost all instances (except for in Swabia and Middle Franconia), which also holds true for a 2003-like weather period (Figure 6.10). However, in case of a 2003-like five-year weather period, AC is also preferred in Swabia.

Overall, farmers in Lower and Middle Franconia were found to be most reluctant to transition away from status quo crop farming. These results show that socioeconomic conditions affect the land-use responses of farmers to regional weather extremes. This involves policy support as well as technological progress. Furthermore, regional differences were found in farmers' willingness to adopt AC and SRC after an extreme weather event. Finally, these results show that prolonged extreme weather periods lead to an increased probability of adopting climate-resilient agroforestry land-use systems in this sample.

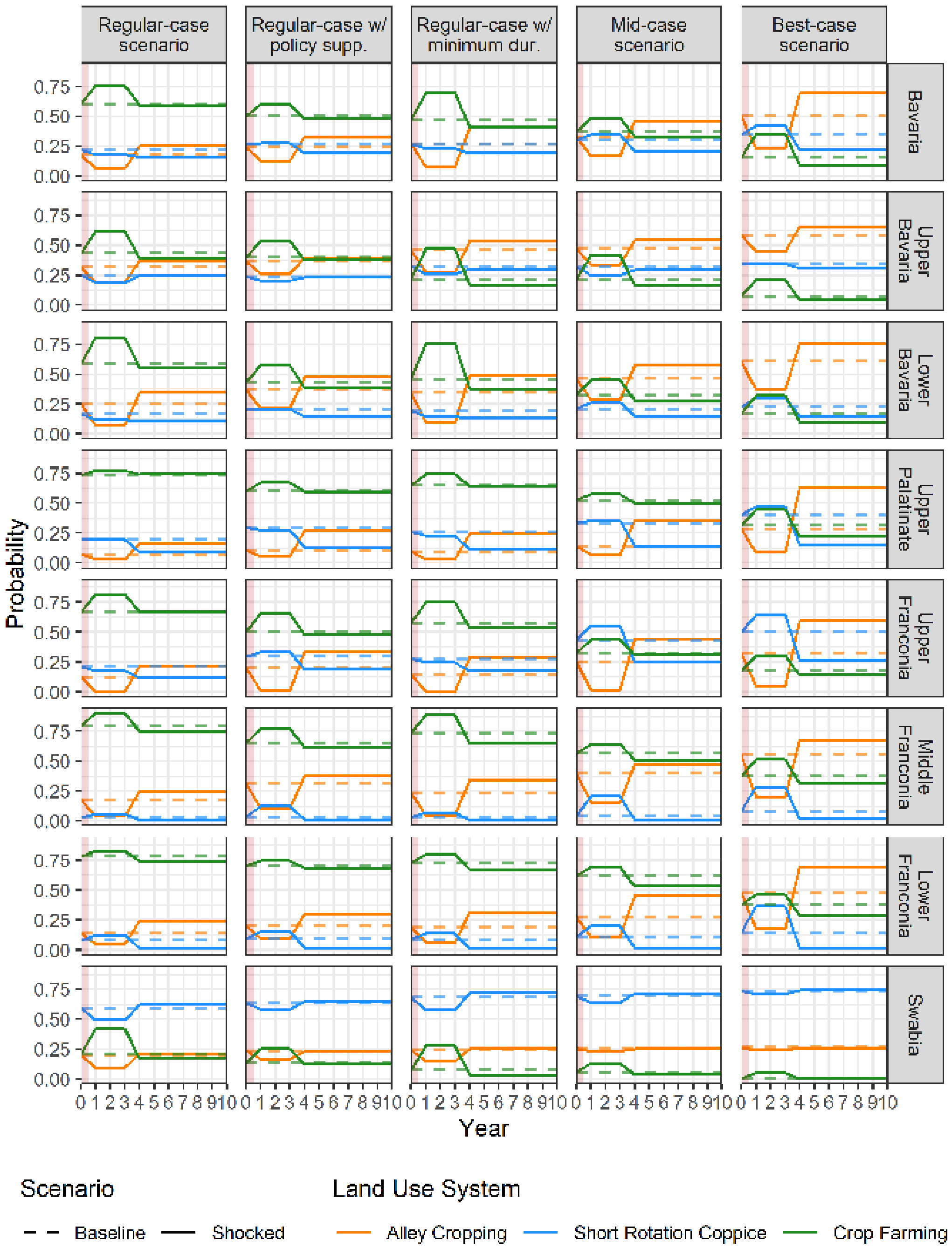


Figure 6.5: Simulated probabilities from a 2018-like extreme weather event lasting one year.

6.6 Discussion

From the result section (6.5), a series of interesting patterns can be observed. In examining farmers' dynamic land-use responses to extreme weather years (Section 6.5.3), one finds characteristic response pathways which occur across farms and regions before reaching a (novel) equilibrium state. More specifically, these pathways can be largely divided into three phases: an absorption phase (during and directly after a shock, in which land-use probabilities move away from the

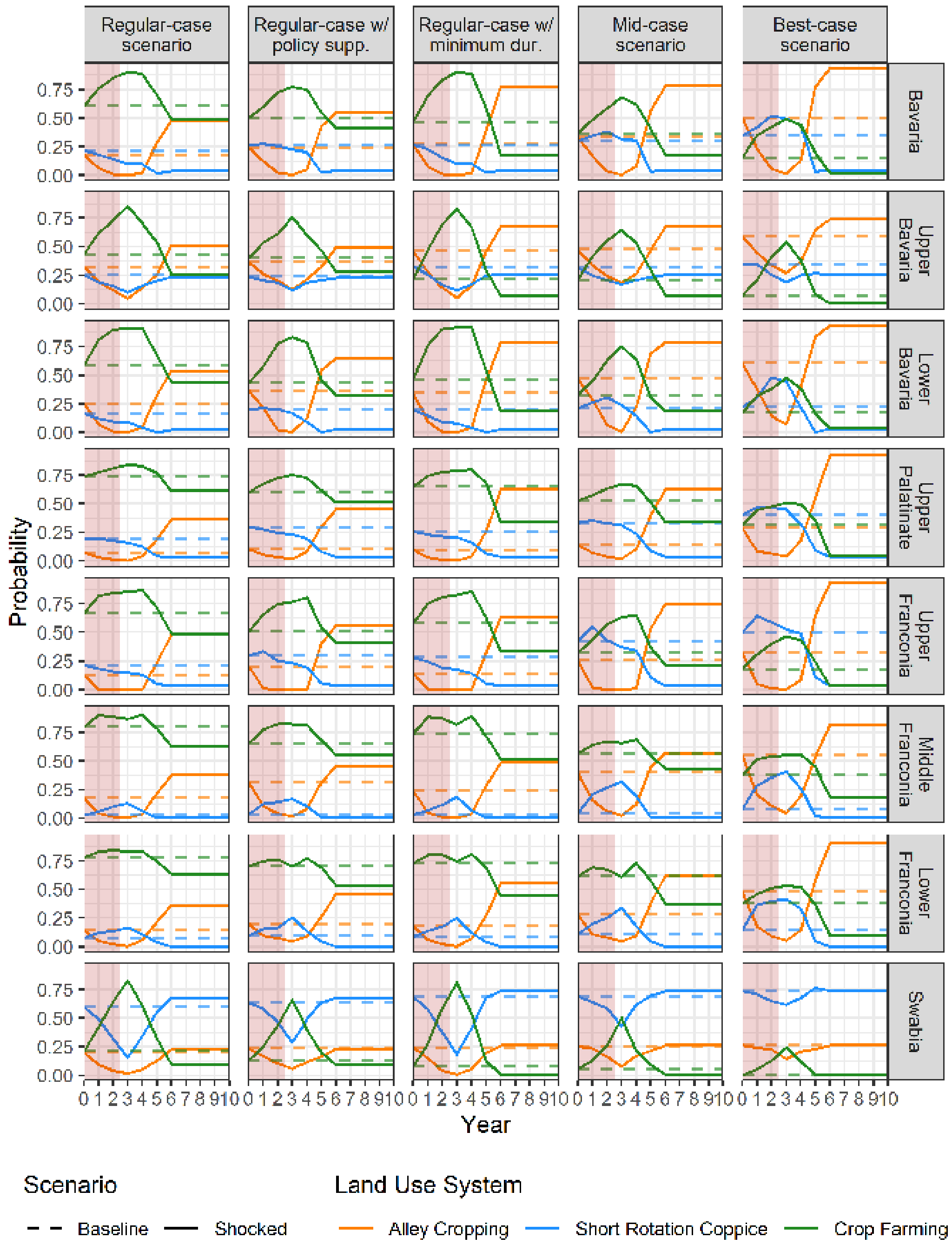


Figure 6.6: Simulated probabilities from a 2018-like extreme weather event lasting three years.

baseline), a recovery phase (in which probabilities return to the initial levels), and an adaptation phase (in which probabilities move away from the initial level toward a (new) equilibrium). These phases reflect important resilience capacities in agricultural systems (Meuwissen et al., 2019; OECD, 2020).

As for the absorption phase, the probability of status quo crop farming was found to increase across all regions encountering a 2018-like weather shock. Since crop farming is also the land-use system with the highest probability in most scenarios

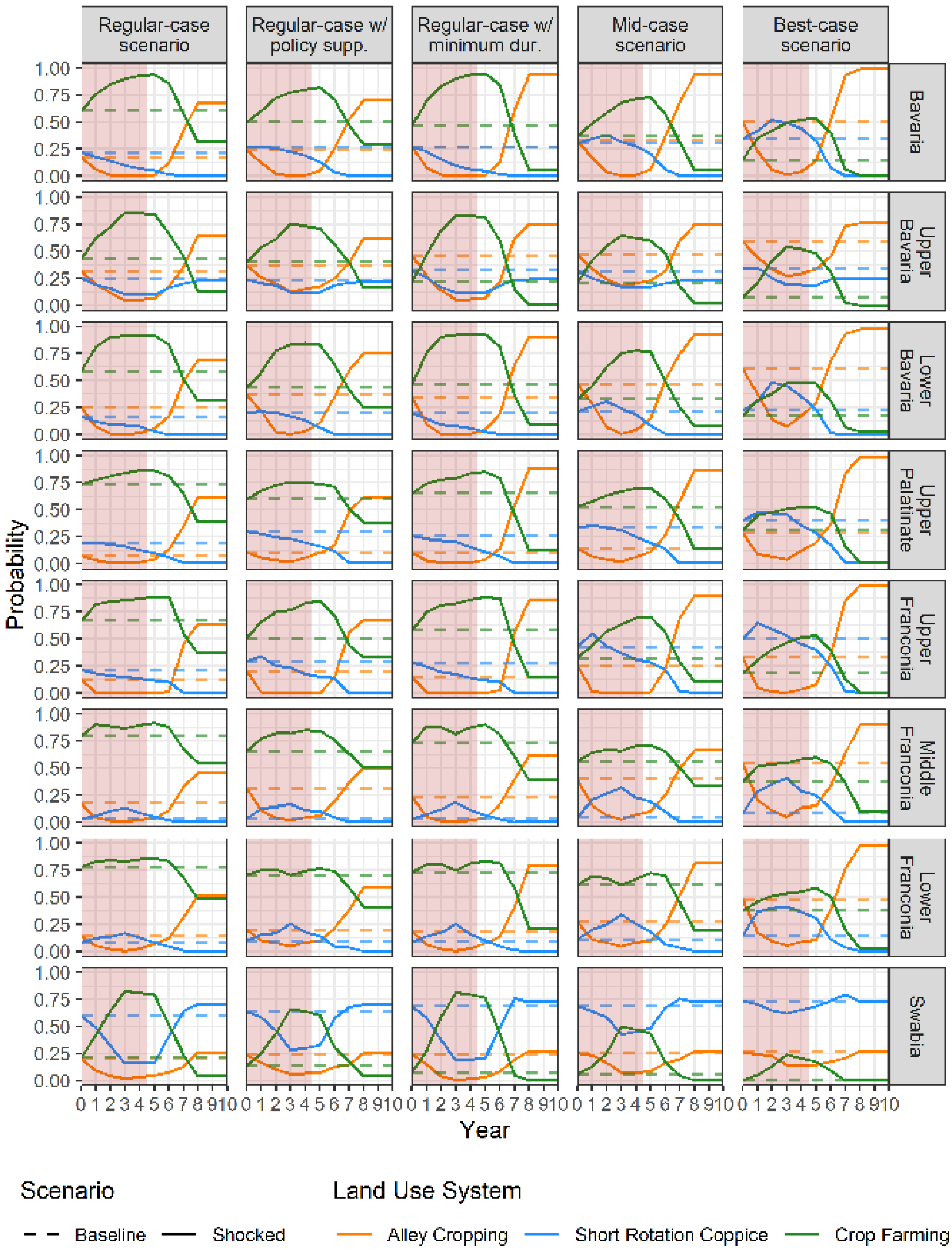


Figure 6.7: Simulated probabilities from a 2018-like extreme weather event lasting five years.

and regions, it can be concluded that farmers adhere more strongly to their status quo in the direct aftermath of a shock when compared to the baseline. This might be surprising at first sight, because one would expect farmers to turn to more weather-robust land uses such as AC and SRC (Ogunbode et al., 2019; Wilson et al., 2020). However, in the short-run, decisional factors are usually rigid, and production structures fixed, and thus farmers' capacity to react is somewhat limited (Girard et al., 2021). Further barriers to transform their land-use directly after a (long-lasting) weather shock might lie in behavioral barriers such as farmers' perceived risk or their perceptions of the benefits and costs associated with more weather-robust land uses (Dessart et al., 2019). Farmers might therefore be prone to only make adjustments within their familiar land-use system (i.e. crop farming). This trend intensifies with the duration of the weather shock.

Depending on the situation, the recovery phase can last between one and five years. Independent of the scenario and region, it can be seen that the sample farms are able to recover from a weather shock in terms of their land-use probabilities (see also Béné et al., 2012; OECD, 2020). This might also be seen as a phase when the extreme weather period has settled and farmers are able to reconsider their initial land use and prepare for adaptive action.

In the adaptive phase, mixed effects regarding farmers' adaptive capacity (Engle, 2011; Smit & Wandel, 2006) are found. AC and SRC both provide comparative advantages over crop farming apart from their relative excellence with respect to climate robustness, but if there is no monetary incentive or technological improvement, farms remain reluctant to transform and adopt these options. Although a certain degree of heterogeneity across shocks and regions was found, this trend was quite stable in our analyses. However, farmers appear to acknowledge the relative excellence of the agroforestry system, because irrespective of the scenario and region, the probability of adopting this system after a weather shock increases in the long-run. Especially, in the case of a very long-lasting extreme weather period (i.e. five years), agroforestry becomes the preferred land-use option.

What is more, our results empirically confirm the conceptual considerations by Meuwissen et al. (2019), i.e. resilience and its capacities are shock- and context-specific.

Next, our results have important implications for policy-makers. First, PES increase farmers' probability of adopting wood-based and agroforestry land-use systems. Therefore, they can be an important lever to promote the cultivation of these climate-robust systems. While PES might be effective in promoting climate-robust land-use systems, they could be more *cost-effective*. From Sec. 6.5.2, we learned that one extra Euro of such payments increases the marginal willingness of most farmers to adopt these system (approx. 67%) by only less than a Euro (median: €0.54). This finding is in line with a series of previous studies finding low cost-effectiveness of PES and agri-environmental schemes (e.g. Bartolini et al., 2021; Chabé-Ferret & Subervie, 2013; Stetter et al., 2022a). However, there is inter and intra-region heterogeneity, e.g. the cost-effectiveness is on average highest in Upper (90%) and lowest in Central Franconia (28%). Accounting for such differences and adjusting the offers of environmental payments for the cultivation of agroforestry at a regional level could significantly increase the cost-effectiveness of such payments (Stetter et al., 2022a; Wünscher et al., 2008).

Another policy-relevant driver of agroforestry adoption is the minimum useful lifetime of the wood-based land use options. Farmers appear to assign a high value to their entrepreneurial flexibility (see also Musshoff, 2012). Rosenqvist & Dawson (2005), Avohou et al. (2011) and Londo et al. (2001) showed that the useful lifetime of wood-based land uses is very important for their economic viability. To better incentivize land-use change, legislators could establish a framework to encourage the development of coppices with reduced minimum useful lifetime but without reduced economic benefits. One way to do this might be the promotion of novel breeding methods, which have shown high innovation potential across several domains (Qaim, 2020).

Furthermore, our analysis adds to a small but increasing body of studies that assesses the link between climate variability and land-use change (Girard et al., 2021). While most of these previous studies focus on established land use types and crops (He & Chen, 2022; Ramsey et al., 2021; Salazar-Espinoza et al., 2015), this study's approach allows the *ex ante* assessment of the potential of novel, not-established land-use types, which could play an important role in the future. What is more, the integration of a choice experiment into the simulation approach allows the evaluation of different scenarios, thus providing a more holistic view on the link between extreme weather and land-use responses.

Finally, this study has some limitations that bear mentioning. For instance, we use cross-sectional weather data for the estimation of the econometric models. This means that farmers' preferences were measured only at one point in time (October 2020). This might be problematic under the assumption that preferences vary temporally (neglecting weather changes, which the model accounts for). However, several studies suggest that preferences are likely to be stable at least in the short- to medium-term (see e.g. Andersen et al., 2008; Dasgupta et al., 2017; Doiron & Yoo, 2017). Another weakness relates to the direct interpretation of the estimated weather coefficients in the RPL model. It is unlikely that any of the weather indicators changes in isolation, i.e. *ceteris paribus* statements are not valid. This is why we refrained from direct interpretation of the estimated weather coefficients and instead focused on the weather simulations, which alleviates this problem to some extent (Ramsey et al., 2021). Additionally, multiple common characteristics of the sample and the underlying Bavarian farmer population indicated reasonable representativeness. This is true for the full sample, but it is not likely the case for the subsample analyses, which is why these results should be interpreted with care regarding their generalizability at the regional level (Pachali et al., 2020). Last, hypothetical bias might be a concern in the presented choice experiment setup. The issue was tackled using cheap talk. There are varying but mostly positive results in the literature with regard to the effectiveness of cheap talk at eliminating hypothetical bias (Liu & Tian, 2021; Murphy et al., 2005; Penn & Hu, 2019). The foregoing notwithstanding, a recent meta-study found that hypothetical bias might not be very problematic in willingness-to-accept settings (Penn & Hu, 2021).

6.7 Summary and concluding remarks

Climate change poses exceptional challenges to farm businesses and the rising number of extreme weather events call for action in terms of climate change adaptation and mitigation. The cultivation of agroforestry and wood-based land-use systems could play a key role in making farms more resilient to climate change. This study analyzes farmers' dynamic WTA such systems in response to extreme weather periods. To this end, random utility theory was integrated with the concept of adaptive weather expectations. Methodologically, a DCE was conducted with farmers in Bavaria, Germany was combined with local weather data. For this analysis, a correlated random parameter model was used, which served as a basis for regional weather simulations following Ramsey et al. (2021).

The results of this study indicate that farms are generally reluctant to adopt agroforestry and SRC compared with crop farming but they are more likely to adopt these options after an extreme weather event in the medium- to long-term. Furthermore, characteristic weather response pathways were found which can be divided into three phases reflecting important resilience capacities, these being absorption, recovery, and adaptation. Additionally, these findings show that policy makers can effectively promote the adoption of agroforestry through PES – although at a relatively low cost-effectiveness – and through fostering technological progress.

Several robustness checks were conducted to assess the plausibility of our model. This study also addresses important limitations concerning the underlying data, its representativeness and the model interpretation. Overall, our results show that farms may be increasingly likely to switch to agroforestry and wood-based systems in response to regional weather extremes.

Finally, we want to outline potential paths for future research. Firstly, it would be worthwhile to assess the statistical uncertainty of the simulations. This could, for instance, be done either by means of a (computationally very expensive) nonparametric bootstrap procedure or by switching to a (hierarchical) Bayesian estimation framework. Furthermore, it would also be interesting to evaluate the appropriateness of this study's approach for other climate change adaptation strategies beyond land-use. Lastly, we would appreciate similar studies in different regions around the world for to get a better overall understanding of the causal links between climate change and land use.

6.8 Appendices

6.8.1 Characterization of the regional districts

Table 6.5: Short summary of the main district characteristics based on Table 6.6.

District	Short characterization
Upper Bavaria 	Large area, highest population density in Bavaria, high livestock density, small and medium sized farms dominate, more than 56% of farms are full-time, for many farms, generational renewal is secured, mostly high altitude, high precipitation, average soil conditions, above-average yields.
Lower Bavaria 	High share of agricultural land, in particular arable land, high land prices, high share of full-time farms, average precipitation and temperature, good soil conditions, above-average yields.
Upper Palatinate 	Highest share of forests in Bavaria, arable land dominates grassland, average farm size structure, below-average land prices, below-average temperature and rainfall, average yields.
Upper Franconia 	High share of rented farmland, high share of large farms, lowest land prices in Bavaria, many part-time farms, below-average temperature and rainfall, low yields.
Central Franconia 	Arable farming is predominant, above-average share of rented farmland, many part-time farms, below-average share of full-time farms, high average temperature, lowest precipitation in Bavaria, below-average yields.
Lower Franconia 	High share of forests, crop farming centered, very low livestock density, highest share of rented land in Bavaria, fewest full-time farms and uncertain generational renewal for many farms, warmest district, below-average rainfall, low yields.
Swabia 	Lowest forest share, high share of grassland-based farming, high livestock density, very few large farms, high share of full-time farms, relatively high altitude, rather low temperatures and much rainfall, region with highest yields.

Table 6.6: Detailed overview of regional district charcteristics primarily based on Bavarian census data from 2010 and 2016.

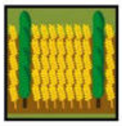
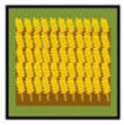
	Bavaria	Upper Bavaria	Lower Bavaria	Upper Palatinate	Upper Franconia	Central Franconia	Lower Franconia	Swabia
Structural conditions								
Total area (mio. ha)	7.1	1.8	1	1	0.7	0.7	0.9	1
Population density (residents/km^2)	186.3	269.3	120.8	114.8	146.9	245.1	154.5	190.7
Share of forest area in total area (%)	36.9	36.2	34.9	43.3	40.6	34	42.2	29.2
Share of agricultural land in total area (%)	44.3	42.8	48.7	40.4	41.3	45.2	41	50.4
Share of arable land in agricultural land (%)	65.5	57.6	74.5	70.1	68.6	70.4	78.5	50.8
Share of grassland in agricultural land (%)	34	42.2	25.3	29.9	31	29.2	19.4	48.8
Livestock density (livestock units/ha agricultural land)	0.9	1.04	0.96	0.93	0.67	0.85	0.37	1.16
Share of rented land (%)	45	36.2	40.3	43.5	50.8	51.1	59	46.9
Size classes (Share of farms in %)								
$<$ 10 ha	22	19.7	23	20.2	25.7	23.3	30.2	18.4
$\geq$ 10 ha $<$ 20 ha	27.2	27.9	27.2	28.8	25.2	27.5	22.6	28.3
$\geq$ 20 ha $<$ 50 ha	29.7	35.3	29.8	28.6	23.6	24.8	22.2	32.2
$\geq$ 50 ha $<$ 100 ha	15.6	13.8	15.8	17.7	16.1	16.8	14	16.7
$\geq$ 100	5.5	3.3	4.2	4.7	9.4	7.6	11	4.4
Land prices (rental price in EUR/ha)	251	272	347	214	148	219	235	282
Full-time farms (%)	51.3	56.9	53.5	48.9	38.9	43.9	38.1	60.3
Generational renewal is secured (% of farms w/ farm manager's age $>$ 45)	37.3	43.8	41.1	39.2	31.3	32.9	28	32.7
Natural conditions								
Elevation (m above sea level)	519.5	630.9	485	489	456.5	411.9	322	646.2
Mean temperature 1960-2020 (°C)	8.09	7.97	8.05	7.89	7.86	8.55	8.76	7.78
Yearly precipitation sum 1960-2020 (mm)	864.4	1046.7	857.8	740.3	796.6	689.6	717.6	988.2
Cation Exchange Capacity[1]	16.5	20.82	13.87	12.11	14.34	15.24	15.37	19.53
pH value of the soil	5.62	5.78	5.48	5.28	5.42	5.71	5.78	5.75
Available water capacity (m^3/m^3)[2]	0.098	0.098	0.095	0.092	0.101	0.099	0.104	0.1
Coarse fragments (%)	14.1	16.2	14.5	13.9	14.7	10.3	12.9	13.5
Yields (dt/ha)								
Wheat	75.1	77.9	80.7	76.3	65.2	70.4	69.7	81.8
Grain maize	104.7	105.3	107.1	103.9	97.6	96.6	97.7	106.2
Grassland	77.1	80	78.5	73.9	69.7	82.6	64.8	88.4

Sources: Bayerisches Landesamt für Statistik (2022), Cornes et al. (2018), Panagos et al. (2012)

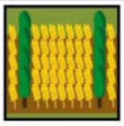
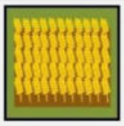
[1] Soil fertility indicator: soil's ability to supply important plant nutrients.

[2] Amount of water that can be stored in the soil and be available for growing crops.

6.8.2 Exemplary choice card (in German)

		Alternative 1: **Kurzumtriebs-plantage** 	Alternative 2: **Agroforst** 	Alternative 3: **Referenz-Ackerbau** 
	Deckungsbeitrag	800€	600€	400€
	Deckungsbeitragsschwankung	+/-30%	+/-30%	+/-15%
	Nutzungsdauer (mind.)	20 Jahre	16 Jahre	3 Jahre
	Agrarumweltzahlungen pro ha und Jahr	0€	100€	0€
	Anerkennung als ökologische Vorrangfläche im Greening?	Ja	Nein	Nein

6.8.3 Description of the DCE alternatives

	Alternative 1: **Short Rotation Coppice**	This alternative relates to the cultivation of fast-growing tree species that are capable of sprouting, such as poplar, willow, locust and others. The trees are planted as a permanent crop on agricultural land and have a harvest cycle of several years. The cultivation is used for material utilization, for example in the paper, pulp and wood-based materials industries. However, the main focus is on generating energy from wood chips.
	Alternative 2: **Agroforestry**	This alternative corresponds to the strip-shaped cultivation of short rotation trees (poplar, willow, robinia and others) in combination with field crops. The distance between the rows of trees is created in such a way that the processing and harvesting of the field crops as well as the energy wood is possible without major restrictions and losses. The proportion of tree strips within the field is approx. 5 or 10%.
	Alternative 3: **Reference Crop Rotation**	This alternative corresponds to a standard crop rotation with the three crops maize, wheat and barley in conventional cultivation. The management requirements correspond to the legal minimum standards.

6.8.4 Description of the DCE attributes

	Attribute	Description
	Margin contribution/ Margin contribution equivalent	The gross margin corresponds to the sum of the revenues minus the variable costs and is the contribution to covering the fixed and overhead costs. In the case of SRC and agroforestry, the relatively high initial investments are offset by only low expenditures for the management of SRC in the following years. Harvest costs and timber revenues are only incurred in years of harvesting activities. By means of dynamic investment calculation (or annuity calculation), the irregular payment flows can be converted into an annual margin contribution equivalent.
	Margin contribution variability	Corresponds to the average fluctuation in the margin contribution (equivalent).
	Minimum useful lifetime	The minimum useful lifetime of the corresponding alternative. For the reference crop rotation alternative, this corresponds to a three-year crop rotation.
	Agro-environmental premium per hectare and year	Annual premium per hectare for the cultivation of the respective land use alternative.
	Recognition as an ecological priority area (greening).	Is it possible to use the area as an ecological priority area?

6.8.5 Weather variable formulation for simulation

One year shock:

$$weath1to3_t = \frac{1}{3} weath_{shock} + \frac{2}{3} weath_{lta} \text{ for } t = 1, 2, 3 \tag{6.9}$$

$$weath1to3_t = weath_{lta} \text{ for } t = 4, \ldots, 10 \tag{6.10}$$

$$weath4to10_t = weath_{lta} \text{ for } t = 1, 2, 3 \tag{6.11}$$

$$weath4to10_t = \frac{1}{7} weath_{shock} + \frac{6}{7} weath_{lta} \text{ for } t = 4, \ldots, 10 \tag{6.12}$$

6.8.6 Weather simulations of a 2003-like extreme weather event

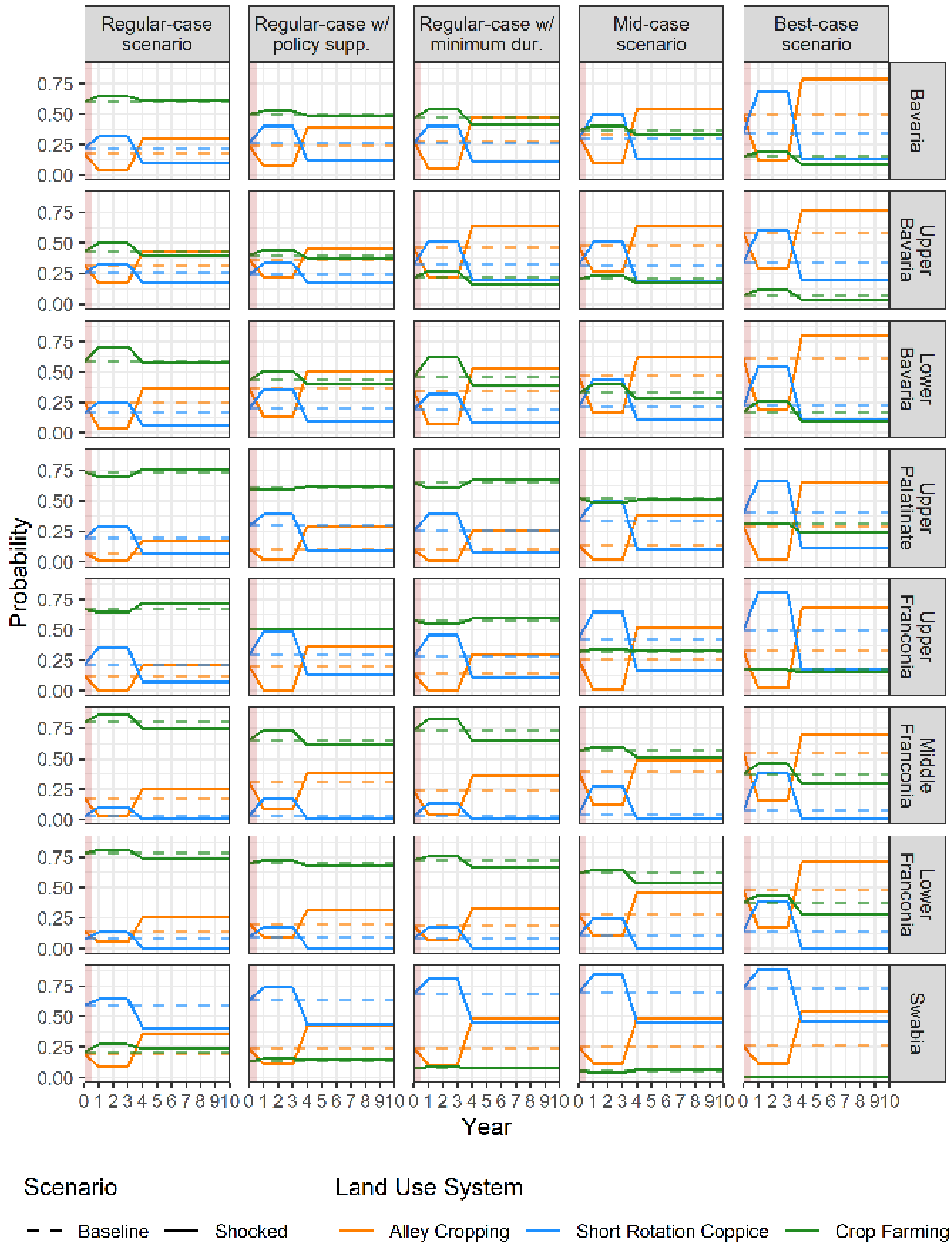


Figure 6.8: Simulated probabilities from a 2003-like extreme weather event lasting one year.

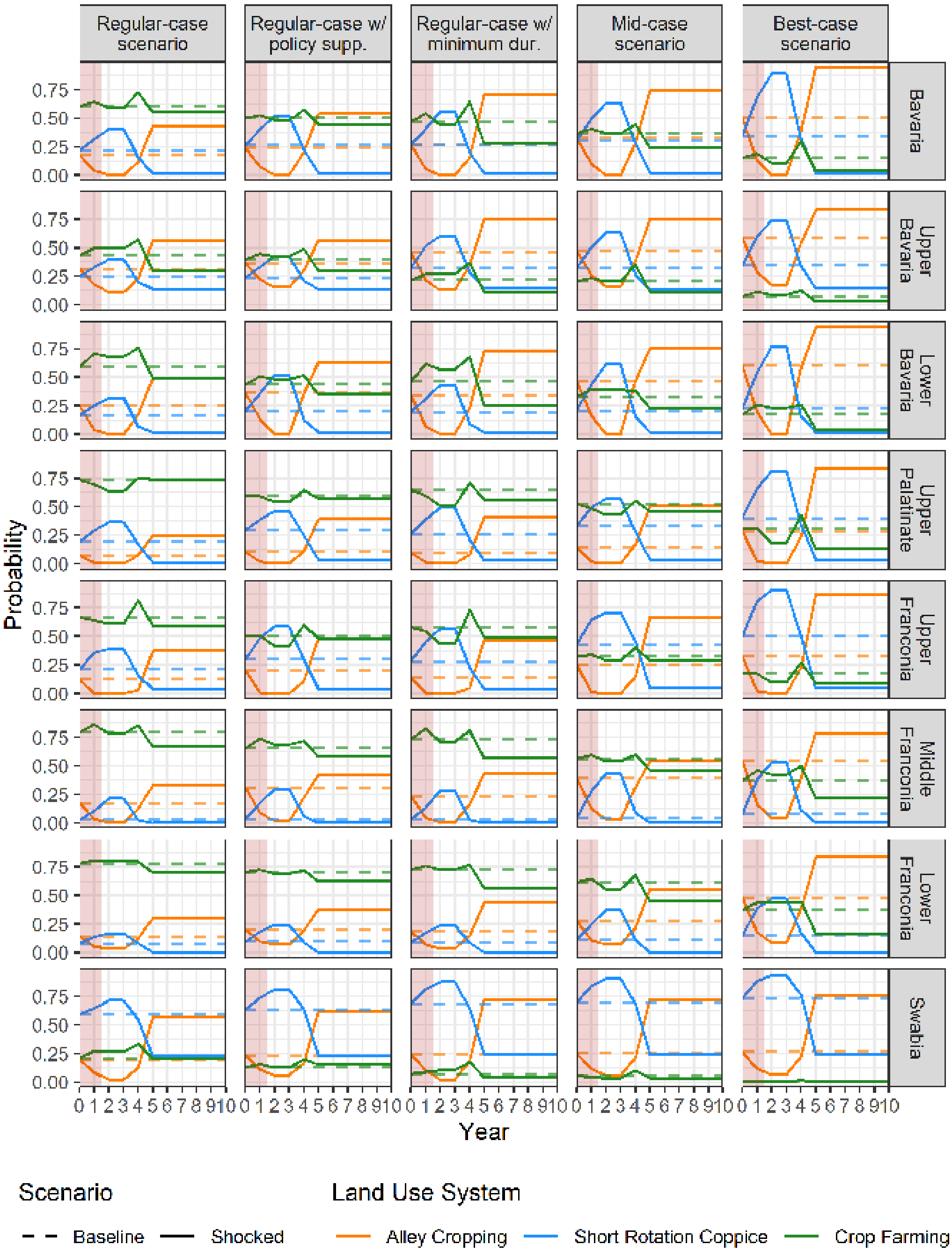


Figure 6.9: Simulated probabilities from a 2003-like extreme weather event lasting three years.

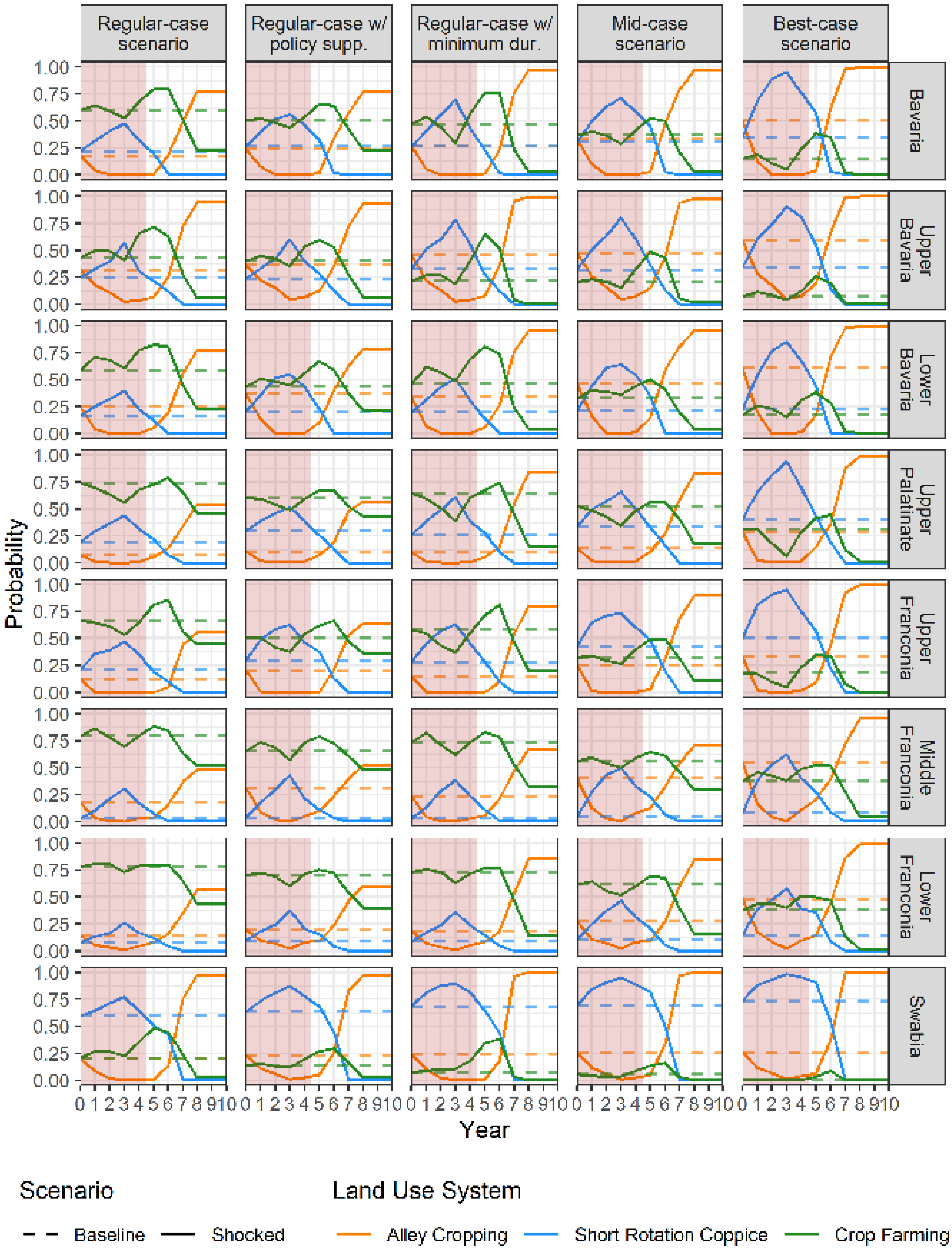


Figure 6.10: Simulated probabilities from a 2003-like extreme weather event lasting five years.

Table 6.7: Estimation results summary.

	MNL	uncor. RPL	cor. RPL
Means			
ASC: SRC	-1.52 (0.22)***	-3.82 (0.57)***	-3.52 (0.60)***
ASC: AC	-0.98 (0.22)***	-2.04 (0.53)***	-0.89 (0.53)$^{\circ}$
Returns	0.00 (0.00)***	-4.87 (0.10)***	-4.80 (0.10)***
Returns variability	-0.01 (0.00)*	-0.03 (0.01)**	-0.05 (0.01)***
Min. useful lifet.	-0.04 (0.01)***	-0.22 (0.03)***	-0.26 (0.03)***
PES	0.00 (0.00)***	0.01 (0.00)***	0.01 (0.00)***
No greening	-0.09 (0.07)	-0.64 (0.17)***	-0.25 (0.22)
Rain 1-3:SRC	0.01 (0.00)*	0.07 (0.01)***	0.10 (0.02)***
Rain 1-3:AC	0.01 (0.00)*	0.02 (0.01)*	0.04 (0.01)***
Rain 4-10:SRC	-0.02 (0.01)*	-0.20 (0.03)***	-0.09 (0.03)***
Rain 4-10:AC	-0.01 (0.01)*	-0.06 (0.02)***	-0.02 (0.02)
Temp. 1-3:SRC	1.17 (0.84)	-5.44 (1.91)**	-9.36 (2.48)***
Temp. 1-3:AC	-1.66 (0.67)*	-14.06 (2.02)***	-1.70 (1.96)
Temp. 4-10:SRC	-1.29 (0.86)	6.05 (1.95)**	8.21 (2.54)**
Temp. 4-10:AC	1.80 (0.69)**	14.25 (2.00)***	1.33 (1.79)
Dry days 1-3:SRC	-0.05 (0.04)	-0.30 (0.13)*	0.69 (0.15)***
Dry days 1-3:AC	-0.03 (0.04)	0.18 (0.10)$^{\circ}$	-0.04 (0.11)
Dry days 4-10:SRC	-0.03 (0.05)	-0.80 (0.18)***	-0.46 (0.20)*
Dry days 4-10:AC	-0.09 (0.05)$^{\circ}$	-0.53 (0.12)***	-0.04 (0.12)
Heavy rain 1-3:SRC	0.03 (0.16)	-0.71 (0.38)$^{\circ}$	-2.87 (0.49)***
Heavy rain 1-3:AC	0.11 (0.13)	-0.27 (0.34)	0.17 (0.35)
Heavy rain 4-10:SRC	0.16 (0.23)	5.48 (0.92)***	3.27 (0.96)***
Heavy rain 4-10:AC	0.02 (0.20)	2.34 (0.56)***	-0.74 (0.56)
Hot days 1-3:SRC	0.01 (0.05)	0.88 (0.18)***	0.70 (0.20)***
Hot days 1-3:AC	0.13 (0.04)**	0.98 (0.16)***	-0.01 (0.15)
Hot days 4-10:SRC	-0.10 (0.11)	-2.01 (0.36)***	-1.75 (0.40)***
Hot days 4-10:AC	-0.18 (0.09)$^{\circ}$	-1.75 (0.30)***	0.30 (0.29)
Standard deviations			
SD Rain 1-3:SRC		0.06 (0.01)***	0.05 (0.00)***
SD Rain 1-3:AC		0.05 (0.00)***	0.02 (0.00)***
SD Rain 4-10:SRC		0.06 (0.00)***	0.05 (0.00)***
SD Rain 4-10:AC		0.05 (0.00)***	0.02 (0.00)***
SD Temp. 1-3:SRC		2.02 (0.57)***	-4.03 (0.58)***
SD Temp. 1-3:AC		2.28 (0.44)***	-0.74 (0.48)
SD Temp. 4-10:SRC		0.02 (0.33)	1.77 (0.46)***
SD Temp. 4-10:AC		-0.81 (0.29)**	-4.44 (0.65)***
SD Dry days 1-3:SRC		0.02 (0.04)	0.09 (0.04)*
SD Dry days 1-3:AC		0.08 (0.02)**	0.00 (0.04)
SD Dry days 4-10:SRC		-0.02 (0.03)	-0.20 (0.04)***
SD Dry days 4-10:AC		0.12 (0.04)**	0.22 (0.04)***
SD Heavy rain 1-3:SRC		0.57 (0.14)***	0.10 (0.18)
SD Heavy rain 1-3:AC		-0.24 (0.15)	-0.37 (0.20)$^{\circ}$
SD Heavy rain 4-10:SRC		0.08 (0.11)	-0.08 (0.17)
SD Heavy rain 4-10:AC		0.11 (0.11)	0.17 (0.11)
SD Hot days 1-3:SRC		0.09 (0.05)	-0.30 (0.06)***
SD Hot days 1-3:AC		-0.25 (0.04)***	-0.28 (0.04)***
SD Hot days 4-10:SRC		0.50 (0.09)***	0.47 (0.11)***
SD Hot days 4-10:AC		-0.78 (0.11)***	-0.57 (0.11)***
SD Returns		1.08 (0.07)***	1.22 (0.69)$^{\circ}$
SD Returns variability		0.07 (0.01)***	0.33 (0.57)
SD Min. useful lifet.		0.25 (0.02)***	1.11 (0.07)***
SD AES		0.01 (0.00)***	0.10 (0.02)***
SD No greening		1.27 (0.20)***	0.28 (0.03)***
SD ASC: SRC		-2.38 (0.32)***	0.01 (0.00)***
SD ASC: AC		-0.33 (0.23)	1.76 (0.28)***
Correlation	-	No	Yes
logLik	-215604	-118413	-115521
Pseudo-R2	0.09	0.50	0.51
AIC	4366.09	2476.27	2460.43
Obs.	2376.00	2376.00	2376.00

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$; $^{\circ}p < 0.1$

6.8.7 Parameter correlation matrix

Table 6.8: Parameter correlation matrix of the RPL model.

	ASC: SRC	ASC: AC	Returns	Returns variability	Min. useful lifet.	PES	No greening
ASC: SRC	0.01	-0.45	0.51	0.29	0.46	0.26	0.57
ASC: AC	-0.45	1.76	0.08	0.11	-0.08	0.09	-0.34
Returns	0.51	0.08	1.22	0.88	0.67	-0.05	0.25
Returns variability	0.29	0.11	0.88	0.33	0.75	-0.43	-0.12
Min. useful lifet.	0.46	-0.08	0.67	0.75	1.11	-0.11	-0.14
PES	0.26	0.09	-0.05	-0.43	-0.11	0.10	0.23
No greening	0.57	-0.34	0.25	-0.12	-0.14	0.23	0.28

6.8.8 Comparison of alternative estimations based on different lag structures

Table 6.9: Comparison of alternative estimations using different lag structures.

	Log Likelihood	McFadden Pseudo R2	Akaike Information Criterion
Selected Model short-term: 1-3 years, long-term: 4-10 years	-1155.21	0.51	2460.43
Specification Alt. 1 short-term: 1 year, long-term: 2-10 years	-1389.99	0.41	2929.97
Specification Alt. 2 short-term: 1 year, long-term: 2-15 years	-1176.99	0.50	2503.98
Specification Alt. 3 short-term: 1 year, long-term: 2-20 years	-1400.24	0.41	2950.47
Specification Alt. 4 short-term: 1-3 year, long-term: 4-15 years	-1238.40	0.48	2626.81
Specification Alt. 5 short-term: 1-3 year, long-term: 4-20 years	-1382.27	0.42	2914.54
Specification Alt. 6 short-term: 1-5 year, long-term: 6-10 years	-1411.71	0.40	2973.42
Specification Alt. 7 short-term: 1-5 year, long-term: 6-15 years	-1398.10	0.41	2946.20
Specification Alt. 8 short-term: 1-5 year, long-term: 6-20 years	-1405.48	0.41	2960.97

Part III

Discussion and Conclusion

7 Summaries and authors' contributions

This thesis aimed at investigating the nexus between agricultural production and environmental change at the micro-level by combining sound microeconomic concepts, state-of-the-art econometric approaches and real-world farm-level data. It has a special emphasis on climate change and encompasses a total of four empirical studies centered around the themes of agricultural production and ecological-economic efficiency, the ecological effectiveness of agri-environmental schemes, and agroforestry adoption in response to extreme weather events. Geographically, all studies focused on Bavaria, a federal state of Germany, one of the core regions of agricultural production within the EU.

7.1 Study I – Greenhouse gas emissions and eco-performance

Given the environmental and climatic implications of farming activities and the need to sustainably increase productivity, Study I focused on the development of a monitoring and evaluation instrument for the ecological-economic performance of farm businesses. It has presented an approach to assess firms' relative ecological (i.e. climatic) damage mitigation potential by building upon and further developing the concept of eco-efficiency (Kuosmanen & Kortelainen, 2005; Orea & Wall, 2017)). A parametric stochastic frontier approach capable of capturing eco-performance dynamics over time was developed. Unlike previous studies on eco-efficiency, this study allowed for a complex functional form to aggregate ecological (climatic) pressures into environmental damage. The resulting *pressure conversion function* described how well ecological pressures translate to economic output. The developed theoretical framework enabled the analysis of eco-performance dynamics and its components – technical change, scale change and eco-efficiency change by means of a generalized Malmquist productivity index.

The empirical application focused on four different farm types in Bavaria for the years 2005 to 2014. A unique combination of various data sources was used to estimate the pressure-generating technology separately for dairy, pig, mixed and crop farms, based on a stochastic frontier model for panel data distinguishing between time-varying and persistent eco-inefficiency.

The main findings of the study were the following. Farms revealed little time-varying eco-inefficiency and rather high levels of persistent inefficiency. Overall, the farms in the sample were quite eco-inefficient. Dairy farms were on average the least eco-inefficient (∼80%) followed by mixed (∼60%), pig (∼55%) and crop farms (∼50%). In terms of eco-performance dynamics, results showed that pig farms revealed the highest annual growth rates. Dairy farms and mixed farms showed less pronounced positive growth. On average, crop farms' eco-performance barely changed between 2005 and 2014.

This work has been published in *Environmental and Resource Economics* (Stetter & Sauer, 2022). The authors contributed to the research article as follows. Christian Stetter designed the conceptual framework of this study. Both authors contributed to the empirical specification. Christian Stetter prepared and cleaned the data, conducted the analysis and wrote the manuscript. Johannes Sauer provided reviewing and editing and contributed with continuous feedback during the entire process.

7.2 Study II – Are intensive farms more emission-efficient?

Study 2 can be seen as a natural extension to the first study. This article evaluated the emission efficiency of distinct technologies in dairy farming, which measures the ability of farms to generate revenue while causing minimal GHG emissions. An eco-efficiency frontier was estimated within a latent-class stochastic frontier framework to identify unobserved heterogeneity in the pressure-generating technology. The results revealed that dairy farms in the case study region could be separated into two distinct technology classes. The two classes could mainly be distinguished by the input intensity of the respective farms in each class. Extensive and intensive farms showed very similar emission efficiency scores when evaluated against their class-specific frontiers. The meta-frontier, which envelops both identified technologies, revealed that extensive farms were overall less emission-efficient than intensive farms, i.e. without losses in economic output, extensive farms could reduce their GHG emissions to 51.5% of current levels when choosing the most efficient technology, compared to 78.9% for intensive farms. Overall, up to 1.7 mio. t CO_2-equivalents (CO_{2eq}) could have been saved in the sample between

2005 and 2014 without reducing economic outcome. Overall, the findings show that technology differences matter, not only with respect to technical efficiency, as suggested by previous research (Alvarez et al., 2012; Martinez Cillero et al., 2018), but also with respect to emission-efficiency. This fact has been largely neglected in previous research on environmental and eco-efficiency. This study showed that the eco-efficiency approach can be employed in a latent-class framework to account for production heterogeneity in environmental efficiency models.

This article has been published in the *Journal of Agricultural and Resource Economics* (Stetter et al., 2022b). The authors contributed to the research article as follows. Christian Stetter and Stefan Wimmer have jointly developed the research idea and both contributed to reviewing and summarizing existing literature. Christian Stetter developed the conceptual framework, constructed the data, estimated the metafrontier and visualized the results. Stefan Wimmer performed the latent-class analysis. Christian Stetter and Stefan Wimmer jointly interpreted the results and wrote the manuscript under the lead of Christian Stetter. Johannes Sauer contributed to the process through valuable suggestions and feedback as well as by reviewing and editing the manuscript.

7.3 Study III – Using machine learning to identify heterogeneous impacts of agri-environment schemes

Study 3 moved away from the relative perspective on the ecological-economic performance of farms and assessed the impact of a policy intervention (agri-environment schemes) on the environmental friendliness of agricultural production at the micro-level by combining economic theory with causal forests, a novel machine learning algorithm based on random forests. The use of this algorithm allowed to evaluate the impact of AES at the farm level and thus delivered valuable information regarding the heterogeneity of the effects of agri-environment measures. The selected research approach presented in this study surpasses many limitations of previous attempts to evaluate the efficacy of AES based on more traditional econometric methods. Conceptually, this study is based on production theory and the potential outcomes framework.

For the empirical case of Southeast Germany, rather small statistically significant effects of AES on land-use diversity were found for approx. 55% of all observations. Regarding fertilizer expenditures per hectare, modest reduction effects for 30% of the sample were found. Desirable effects on pesticide expenditures could be found for 7% of the sample. In terms of GHG emissions, mostly insignificant or adverse effects were found. The findings of the study point toward the direction that treatment effects of agri-environment measures on important environmental indicators were rather small during the 2014-2020 CAP period.

Based on these results, spatial patterns of environmental subsidy payments as well as important drivers of heterogeneous treatment effects could be explored. A large share of desired effects in at least one environmental dimension was detected in almost all counties. Using Shapley values to predict the contribution of the four dimensions location, farm type, yield potential and farm size, it could confirmed that targeting of agri-environment payments could potentially improve environmental efficacy for all environmental indicators used in this study. Targeting farms in terms of location, farm size, and yield potential by nudging for example can result in more efficient usage of environmental subsidies while targeting schemes according to different farm types did not seem to drive subsidy effectiveness. Finally, a battery of sensitivity tests was used to assess the robustness of the results in various settings.

This article has been published in the *Journal of Agricultural and Resource Economics* (Stetter et al., 2022a). The authors contributed to the research article as follows. Christian Stetter and Philipp Mennig have both contributed to reviewing and summarizing existing literature on AES. Christian Stetter developed the conceptual and empirical framework, constructed the data, trained the model and conducted all analyses. All authors interpreted the results. Christian Stetter and Philipp Mennig jointly wrote the manuscript under the lead of Christian Stetter and wrote the manuscript under the lead of Christian Stetter. Johannes Sauer contributed to the process through valuable suggestions and feedback as well as by reviewing and editing the manuscript.

7.4 Study IV – Agroforestry adoption and weather extremes

The cultivation of agroforestry and wood-based land-use systems could play a key role in making farms more resilient to climate change. This study analyzed farmers' dynamic willingness (i.e. probability) to adopt such systems in response to extreme weather periods. To this end, random utility theory was integrated with the concept of adaptive weather expectations. Methodologically, a DCE was conducted with farmers and combined with local weather data. For this analysis, a correlated random parameter model was used, which served as a basis for regional weather simulations following Ramsey et al. (2021).

The results of this study indicated that farms were generally reluctant to adopt agroforestry and SRC compared with crop farming but they are more likely to adopt these options after an extreme weather event in the medium- to long-term. Furthermore, characteristic weather response pathways were found, which could be divided into three phases reflecting important resilience capacities, these being absorption, recovery, and adaptation. Additionally, the findings showed that policy makers could effectively promote the adoption of agroforestry through PES – although at a relatively low cost-effectiveness – and through fostering technological progress. Several robustness checks were conducted to assess the plausibility of the estimation model. This study also addressed important limitations concerning the underlying data, its representativeness and the model interpretation. Overall, the results showed that farms might be increasingly likely to switch to agroforestry and wood-based systems in response to regional weather extremes.

This work is currently under review at *Environmental & Resource Economics*. The authors contributed to the research article as follows. Christian Stetter developed the conceptual and empirical framework of the study. Christian Stetter prepared and conducted the discrete choice experiment, retrieved the weather data and joined the datasets. Christian Stetter conducted the analyses and wrote the manuscript. Johannes Sauer contributed to the process through valuable suggestions and feedback as well as by reviewing the manuscript.

8 Discussion and policy implications

The focus of this chapter lies on the synthesis of the results and conclusions drawn from the empirical studies in Part II of this thesis. It further discusses important, cross-cutting themes in this context. Moreover, multiple policy implications are explored. Finally, it highlights further research directions, which can be derived from this thesis.

8.1 General reflections on the economic-ecological performance of farms

Having analyzed the relationship between agricultural production (reflected by economic returns) and environmental degradation (reflected by GHG emissions), the first two studies of this thesis showed that environmental protection (i.e. climate change mitigation) and farming do not necessarily have to be mutually exclusive. This finding was robust across various farm types and technologies. Farms can drastically decrease their climate change impact without risking economic viability by improving farm management practices, i.e. reducing eco-inefficiencies. From an eco-efficiency perspective, there is a lot of potential for farms to improve in this regard. This finding is in line with the recent eco-efficiency literature and has been found for other environmental stresses as well, such as fertilizer and pesticide damages (e.g. Bonfiglio et al., 2017), water usage (e.g. Song & Chen, 2019), or biodiversity loss (Beltrán-Esteve et al., 2014). This means there is ample room for a synergistic relationship between reducing farms' environmental impact and being economically successful up to an (eco-)efficient frontier.

This raises two central questions. First, why are farmers eco-inefficient? Second, what is the economic-ecological relationship on the eco-efficient frontier? As for the first question, this thesis does not have a definite answer. However, from a farmer's perspective, it is not unlikely that a certain degree of eco-inefficiency is rational, i.e. they actively choose not to be eco-efficient. Usually in classical production economics, technical inefficiency is regarded as waste in the utilization of production factors (Bogetoft & Hougaard, 2003; Hansson et al., 2020). However, eco-efficiency is conceptually different from technical efficiency and gives equal weight to ecological and economic objectives. In reality, farms might predominantly aim at optimizing economic performance, which could potentially (but not necessarily) cause eco-inefficiencies for the sake of higher economic efficiency. Therefore, if the aim is to reduce eco-inefficiency, mechanisms might be necessary that provide incentives to farmers to improve their ecological performance, e.g. farm management practices that provide both economic and ecological merits (e.g. higher milk yield as shown in Study 2).

As for the second question, it could potentially be the case that reductions of environmental pressures beyond the eco-efficient production level are inevitable in order not to irreversibly surpass earth's planetary boundaries (Kuosmanen & Kortelainen, 2005). For this case, Sauer & Wossink (2013b) have empirically shown that there exist three relationships between agricultural production and non-marketed ecosystem services generation: complementary, supplementary and competitive. For the case of arable farms in the UK, they found a competitive relationship for many instances, i.e. decreasing the environmental impact means a net loss in agricultural output. Overall, the results from the literature on this relationship have been mixed and context-specific (Rosa-Schleich et al., 2019), e.g. Kragt & Robertson (2014) found both win-lose and win-win trade-offs for broadacre farming in Western Australia, and Galdeano-Gómez et al. (2017) found a mostly synergistic relationship for fruit and vegetable farms in South Spain.

What is more, Study 2 laid the focus on the GHG mitigation potential in dairy farming. Dairy farming is likely to play a key role in the next decades when it comes to GHG emission reduction and tackling climate change. This is because most GHG emissions on dairy farms come from methane, which has a high global warming potential but only a short atmospheric lifetime (Pérez-Domínguez et al., 2021). This is why its reduction offers the possibility to mitigate climate change efficiently, especially in the shorter run (Saunois et al., 2016).

Furthermore, a pervasive assumption in the public debate on agriculture in Europe is that extensive production systems are inherently more environmentally-friendly than their intensive counterparts. However, the results from Study 2 show that GHG emissions translate more efficiently to economic returns for intensive dairy farms in Bavaria. Again, this implies a relative perspective on the environment-productivity nexus. The relative excellence of intensive dairy farms in this regard could come from the fact that they are likely to have more control over nutrient flows and animal behavior through better infrastructure and management skills (Alvarez & del Corral, 2010). Burney et al. (2010) found a similar positive effect of agricultural intensification in terms of GHG mitigation for crop farming on a global scale.

However, one should not jump to conclusions and promote the intensification of agriculture as an imperative for environmentally-friendly production. The literature shows that the intensification of agriculture has also had negative impacts on various environmental domains not assessed in this thesis such as soil health (Kopittke et al., 2019), biodiversity (Ewald et al., 2015; Gossner et al., 2016; Ramos et al., 2018; Seibold et al., 2019), or water quality (Scanlon et al., 2007).

Hence, in terms of the production-environment nexus, there might also be important trade-offs among environmental objectives themselves. The identification of such trade-offs (and synergies) beyond climatic stresses would be an important addition to the results of this thesis. However, in the European context it has been difficult to obtain the necessary data for such analyses (see below).

8.2 The role of farm system dynamics

Another important cross-cutting theme that has been considered in this thesis is the fact that farming systems are dynamic by nature and evolve over time (Dillon, 1992; Meuwissen et al., 2019), which also has consequences for the production-environment nexus and agroforestry adoption.

First, Study 1 distinguished between time-varying (short-term) and persistent (long-term) GHG inefficiency. These two inefficiency measures relate to different managerial dimensions covering different time horizons. Time-varying inefficiency reflects short-run rigidities, temporary managerial and behavioral problems that are potentially solvable in the short-run (Addo & Salhofer, 2022), whereas time-invariant inefficiency reflects the "[...] presence of structural problems in the organization of the production process of a firm or the presence of systematic shortfalls in managerial capabilities" (Filippini & Greene, 2016, p.187). Study 1 finds that, independent of the farm type, permanent eco-efficiency is noticeably lower than time-varying efficiency. From that, it becomes apparent that efficiently reconciling economic success and GHG mitigation requires structural, longer-term adjustments, which has important consequences in that such adjustments usually take time and cannot be made in the short-run (Filippini & Greene, 2016).

A similar result was also found in Study 2 in that structural change in the form of a technology switch for extensive farms could lead to higher GHG mitigation at constant economic returns. Furthermore, Study 4 is indicative of the fact that short-run rigidities keep farmers from making optimal decisions in terms of agroforestry adoption in the aftermath of an extreme weather year. Farm system dynamics appear to prevent farmers from making short-term adjustments to improve climate change mitigation and adaptation but rather require adaptive and transformative capacities (Spiegel et al., 2020). Such further-reaching adjustments depend upon farmers' management abilities and cognitive social capital and might only be carried out after certain thresholds of a key decision variables have been exceeded (Sinclair et al., 2014).

Next, the fact that farming at the nexus of agricultural production and environmental change is not static can also be seen by farms' eco-performance evolution over time (Study 1). Within the analyzed time horizon, GHG emission performance across farm types is subject to yearly fluctuations of up to more than 10% (see Table 3.8 and Figure 3.3), which highlights the dynamic nature of this process, which has often been neglected in the literature (e.g. Beltrán-Esteve et al., 2014; Godoy-Durán et al., 2017; Martinsson & Hansson, 2021). The importance of the time dimension has also become apparent in the simulation of agroforestry adoption patterns (Study 4). Farmers react dynamically to a weather shock in the years following such a shock. This is in line with recent literature on resilience in agriculture (Meuwissen et al., 2019).

8.3 Production heterogeneity and context specificity

Another recurring theme across all studies is the fact that the production-environment nexus is not homogeneous but context-specific. This thesis assessed differences across and within farm types regarding eco-efficiency, regional differences regarding the adoption of agroforestry as well as farm-to-farm differences for the environmental effectiveness of AES (see Table 8.1 for an overview).

Tsionas (2002) argue that the pervasive assumption in the literature that production is the same for all farms is quite unrealistic and an over-simplification. The production possibilities of farms are usually bound to specific technologies, which cannot be easily switched (e.g. crop farming vs. livestock farming, or grassland vs. arable farming), which makes it important to distinguish between farm types (Study 1), technological (Study 2), and/or regional differences (see Study 4). Study 3 went even a step further by explicitly defining the farm-specific context through a comprehensive set of approx. 130 covariates.

Table 8.1: Level of farm & production heterogeneity as considered in Part II.

Study/ chapter	Research focus	Level of heterogeneity
Empirical study 1/ chapter 3	Eco-efficiency and eco-performance w.r.t. GHG emissions	Farm type differences (dairy, pig, crop, and mixed farms) *(ex ante)*
Empirical study 2/ chapter 4	Eco-efficiency w.r.t. GHG emissions	Production intensity (intensive & extensive dairy farms) *(data-driven)*
Empirical study 3/ chapter 5	Heterogeneous effect of agri-environmental schemes on the environmental performance of farms	130 contextual variables to account for individual farming context *(data-driven)*
Empirical study 4/ chapter 6	Agroforestry and wood-based land-use systems adoption in response to extreme weather	Regional differences (district level) *(ex ante)*

The fact that the results of all four studies vary across farms and groups of farms confirms the presupposition that production heterogeneity is an important factor when it comes to the analysis of the production-environment nexus. This finding is in line with the recent agricultural production economics and impact assessment literature (see e.g. Ait Sidhoum et al., 2022; Baráth et al., 2020; Njuki et al., 2019; Sauer & Moreddu, 2020; Sauer & Wossink, 2013a) and has important consequences for policy-makers and agricultural stakeholders (see below).

A critical point worth discussing in this context is the potential conflict between context-specificity and generalizability. In this thesis, Bavaria was intentionally chosen as a case study. As shown in the introduction of this thesis, it might be well-representative of many European regions. Nevertheless, extrapolating results to other regions should be done with care and should further be evaluated by local experts. Furthermore, on a lower level, Study 4 collects primary data of Bavarian farmers, which reflect the Bavarian farmers population reasonably well (see Sec. 6.5.1). While this holds for the full sample, it is not likely the case for the subsample analyses, which is why results should be interpreted with care regarding their generalizability at the regional level (Pachali et al., 2020). This problem is likely reduced in the case of the FADN dataset used in Studies 1-3 as it provides representative data according to region, economic size and type of farming (European Commission, 2022b). Nevertheless, context-specific, heterogeneous findings should be interpreted with care in terms of their generalizability to the underlying populations.

8.4 Methodological and conceptual contributions

Beside the empirical insights into the agricultural production-environment nexus, this thesis makes several methodological and conceptual contributions to the agricultural economics toolset.

To begin with, Study 1 is the first study that accounts for both persistent and time-varying eco-efficiency. Although there have been multiple methodological and conceptual advancements since the introduction of the concept by Kuosmanen & Kortelainen (2005, see also Sec. 2.1.3), none of them accounted for the fact that eco-efficiency can be separated into two important components following recent developments in the productivity analysis literature (Filippini & Greene, 2016; Kumbhakar et al., 2014). Study 1 and the discussion on farm dynamics demonstrated the added value by using this approach.

Furthermore, based on the parametric eco-efficiency concept developed in Study 1, Study 2 demonstrates how eco-efficiency and latent-class analysis can be combined to take account of technological differences among groups of farms. If such technology differences are ignored, it might be the case that they are wrongly labeled as inefficiencies leading to biased results (Orea & Kumbhakar, 2004). There are already eco-efficiency studies comparing different technology groups. However, they are commonly chosen *a priori* based on observed differences, e.g., Beltrán-Esteve et al. (2014) in their study on rain-fed olive farming system in Spain demonstrate the relevance of technology differences between traditional mountain and plain groves with respect to eco-efficiency. This thesis followed the strain of literature that propagates the use of latent-class modeling because this approach is able to account for latent (unobserved) technology differences, which would otherwise not be possible (Alvarez & del Corral, 2010; Martinez Cillero et al., 2019; Orea & Kumbhakar, 2004).

Furthermore, the research approach presented in Study 4 goes well-beyond the usual application of discrete choice experiments through the combination of preference and meteorological data, regression analysis and simulation. There are a few similar studies in different contexts, e.g. Dundas & von Haefen (2020) assessed the effects of weather on recreational fishing demand and adaptation. Motoaki & Daziano (2015) analyzed the weather effects on cycling demand, Hashida & Lewis (2022) focused on the climate change impact on forest management in the Pacific states of the U.S. However, none of these studies accounted for the dynamic nature of adaptation decisions and neglected the multitude of potential future socioeconomic scenarios as it is done in this thesis.

Finally, Study 3 constitutes the first study to demonstrate the applicability of causal machine learning to the environmental impact assessment of AES in the EU. The utilized causal forest algorithm allows to better capture the high degree of complexity, nonlinearity as well context-specificity of the AES-environment nexus in the realm of agricultural production. The approach overcomes several limitations of standard econometric approaches, including restrictive functional forms and the limited ability to deal with a large number of explanatory variables (Storm et al., 2020). As is often the case in machine learning (ML), the increased model complexity of the causal forest comes at the expense of interpretability (James et al., 2021). This problem was tackled by means of model-agnostic Shapley values, which were further used to draw conclusions with respect to scheme targeting. A pervasive problem with such flexible machine learning approaches is that they allow researchers to include a myriad of predictor variables in their models, which are prone to lead to bias structures in cause-effect relationships, e.g. by including bad control variables (Cinelli et al., 2020; Hünermund et al., 2021). Thus, if one is interested in causal inference rather than pure outcome prediction, it is absolutely necessary to set up a credible identification strategy based on a comprehensive theoretical model (Pearl, 2018). Modern machine learning algorithms do not free the applied researcher from conducting this first stage, which is why Study 3 lays special emphasis on a solid conceptual framework.

Finally, it is important to note that the use of Shapley values as model-agnostic interpretation tool refer to the modeled relationship between AES and the environmental indicators and not the ground truth, which is not the same (Lipton, 2018). This is a major epistemological difference from more traditional statistical methods. Hence, the results on CATE drivers in Study 3 rely on the assumption that the empirical model approximates the causal mechanisms of the true relationship sufficiently well (Páez, 2019). Bearing this in mind is particularly important if one was to include an excess set of covariates (see above) that are not causally related to the outcome variable, which would lead to Shapley values explaining a spurious relationship. For the assumption that the estimated model reflects the true relationship, beside having a reliable identification strategy, it is important to conduct various robustness checks to see if the estimated relationship is indeed stable across multiple configurations of the model as done in Study 3 of this thesis.

8.5 The importance of data availability, sources, and processing

The availability of empirical data plays a key role in this thesis. Studies 1 – 3 use primarily farm accountancy data, which are collected for the Farm Accountancy Data Network of the EU (FADN). The underlying farm-level GHG indicator additionally uses a series of secondary data and merge these with the FADN dataset.

The FADN monitors commercial farms' income and business activities and it is the only source of microeconomic farm data based on harmonized bookkeeping principles (European Commission, 2022b). This dataset is a very important source for myriad studies on agricultural production in the context of the EU. However, the dataset is not explicitly built for the evaluation of environmental dimensions of farming such as environmental sustainability, climate change adaptation and mitigation, or resource efficiency. The data fusion approach used to develop a farm-level GHG emission approximation (Study 1) can be seen as an important step to overcome the lack of environmental information in the FADN dataset, but it cannot divert attention from the fact that better farm-level environmental data is needed.

Data limitations are also relevant for Study 3, which suffers mainly from two shortcomings of the dataset. First, the environmental indicators used as response variables are only coarse approximations of the underlying environmental domains. Except for the case of GHG emissions, the utilized indicators do not measure direct environmental impacts like water pollution or soil degradation. However, more accurate environmental indicators are very important when one wishes to assess the environmental effects of the EU's agricultural policies. Second, to draw more precise policy conclusions, it would be necessary to provide more detailed data on individual AESs. By treating AES participation as a binary variable, a lot of important information is lost as the multitude of different AES, which might be targeted toward different environmental services is not properly reflected in the model. Thereby, an assessment of the environmental performance of individual AES should be in the policy-makers' own interest as this could lead to more specific, evidence-based policy recommendations.

Given these current limitations of the FADN dataset, the European Commission has realized the need to further develop this data source against the background of the European Green Deal and its Farm to Fork strategy (Vrolijk & Poppe, 2021). It has decided to convert the FADN to the Farm Sustainability Data Network (FSDN). From a researcher's perspective this can be regarded as huge leap forward to create further sound and evidence-based knowledge regarding farms' economic and environmental performance. First steps have already been undertaken in the past years, e.g. the inclusion of detailed information on fertilizer and farm chemicals usage. Further extending the FADN dataset by environmental and social dimensions implies several difficulties, however, such as the high cost of data collection, farmers' limited willingness to provide the respective information, increased administrative burden, and a potentially limited quality of data (Vrolijk & Poppe, 2021). The European Commission is currently in the process of setting up a roadmap for the conversion of the data network and has received a first round of feedback from various stakeholders (European Commission, 2022a). It remains to be seen if the Commission will be able to solve the above-mentioned issues and set up an FSDN that provides usable information for better sustainability assessments at the farm level.

An alternative to retrieving environmental information through survey responses are geospatial data, which have become increasingly available in recent years and are heavily driven by the advancement of satellite remote sensing (Chi et al., 2016; Kamilaris et al., 2017). In Europe, big geospatial data is available on a fine temporal and spatial scale, which is particularly true for weather and climate data (Kamilaris et al., 2017). Other environmental domains include, e.g., soil properties (Panagos et al., 2015, 2012), invasive species (Katsanevakis et al., 2012), land use change (Winkler et al., 2021), or air pollution (Cui et al., 2021). Acknowledging the additional value of geospatial data, Study 4 combines an economic experiment with gridded meteorological data, thus allowing to assess the relationship between weather (expectations) and farmers' land use preferences. The integration of economic and biophysical data has increased considerably during the past years and has provided novel insights, e.g., in the fields of agricultural insurance

(Vroege et al., 2021), productivity analysis (Ortiz-Bobea et al., 2021), or land use (Ramsey et al., 2021).

While the combination of economic and geospatial data has large potential, it is also associated with various difficulties. For instance, merging economic and spatial data requires knowledge of the exact location of farms and/or fields. Researchers usually do not have this knowledge in the context of the FADN for data protection reasons. Thus, only analyses at an aggregate (e.g. regional) level are possible, which itself leads to imprecisions and spatial mismatches between economic and geospatial data. What is more, especially remotely-sensed data are often measured at a fine temporal resolution such as hours or days, while farm economic data is usually collected annually, which can cause the problem of finding the appropriate joint level of resolution (Arbia, 1989).

Finally, as mentioned above and demonstrated in Study 3, novel ML-based methods can bring about additional value to the analysis of farm production processes and the environment. However, they are usually data-intensive, i.e. they require a large sample of observations. Economic and survey data generated through interviews can often be on the lower end regarding the minimal sample size necessary for ML algorithms to work (which is especially true for neural networks). Hence, in the face of limited data availability, such methods should not be regarded as a magic bullet, but rather as a potent extension to the applied researcher's toolset applicable in certain settings .

8.6 Policy implications

As suggested before, the results presented in this thesis have several policy implications. First, the potential of AES to promote environmentally-friendly and eco-efficient farming practices have been mentioned multiple times throughout this thesis and were the primary research subject of Study 3. Table 8.2 summarizes the effects of AES on different (economic-)environmental indicators.

While Study 1 did not explicitly look at the effect of AES, Study 2 found a negative association between AES and eco-efficiency. This is somewhat in contrast to previous work, which found mainly a positive relationship (Bonfiglio et al., 2017; Gadanakis et al., 2015; Picazo-Tadeo et al., 2011). A potential reason for this could be the fact that the other studies did not explicitly consider GHG emissions (i.e. climate impacts). One conclusion of this could be that the effectiveness of

Table 8.2: Overview of the effects of agri-environmental schemes presented in Part II.

Study/ chapter	Indicator	AES effect*
Empirical study 1/ chapter 3	Eco-efficiency w.r.t. GHG emissions for dairy, pig, crop, and mixed farms	*Not evaluated*
Empirical study 2/ chapter 4	Eco-efficiency w.r.t. GHG emissions for intensive and extensive dairy farms	−
Empirical study 3/ chapter 5	GHG emissions	0/−
	Fertilizer expenditures per ha	0/+
	Pesticide expenditures per ha	0/+
	Land use diversity	+/0
Empirical study 4/ chapter 6	Agroforestry and wood-based land-use systems adoption	+

* + Desired effect; − Undesired effect; 0 No effect.

AES in terms of eco-efficiency depends on the underlying environmental indicators, e.g. climate change had not been a priority in CAP pillar 2 schemes, while other pressures enjoyed more attention (e.g. nutrient runoff and diversification, ART, 2019). The thematic coverage of AES was extended to climate objectives only in 2009 following the CAP Health Check, and only as of 2014, AESs have been referred to as 'agri-environment-climate schemes', emphasizing current and future climate change mitigation and adaptation efforts (see also Hasler et al., 2022). Notwithstanding, the practical implementation of climate objectives have been lacking behind ever since (European Court of Auditors, 2021).

Having recognized the urgent need for action in terms of climate change and having set the objectives of reducing GHG emissions by at least 55% by 2030 and becoming climate neutral by 2050 (Hasler et al., 2022), agricultural legislators should more strongly focus on effective AES design in terms of GHG mitigation. This could be achieved by promoting the adoption of existing climate-friendly technologies.

One such climate-friendly farming practice is the cultivation of agroforestry systems, which was analyzed in Study 4. Following the current proposal on the implementation of the post-2020 CAP, agroforestry systems will be financially supported by way of the newly created eco-scheme framework (Deutscher Bundestag, 2021). In Germany, it is planned that farmers receive €60 per hectare with regard to the tree-covered area. This would mean in a system with 10% tree-covered area, €6 per ha (Landwirtschaftskammer Niedersachsen, 2022). Given our estimation results of a negative willingness to cultivate alley cropping of €123 on average for Bavaria and a positive marginal WTA of €0.53 for PES, it would require financial support of at least €232 per ha for agroforestry to be adopted (with regard to total area, incl. crops and trees). Although this support sum varies across regional districts and farms (see Study 4), the current proposal will very likely not lead to a wide adoption of agroforestry among farmers in Germany and Bavaria.

This reveals another problem of the currently planned implementation of the CAP reform. Production heterogeneity and context specificity are not explicitly taken into consideration. This is in contrast to one of the main conclusions of Study 3, namely that it is essential to take account of these aspects as they can be used for targeting, which can ultimately lead to greater environmental effectiveness of AES. Based on context information, farms with high predicted participation effects could be encouraged to participate in AES through different approaches, such as paying a collective cohort bonus, reducing transaction costs, linking payments amounts to site conditions, introducing spatially-coordinated auctions for conservation contracts or other incentive payments (Del Rossi et al., 2021; Ferraro, 2008; Kuhfuss et al., 2016; Pelosi et al., 2010).

Furthermore, given the low effectiveness of AES across indicators (Study 3), European legislators might have to fundamentally reconsider and revise agri-environment payment schemes. The literature has made several suggestions including, e.g., result-based payments (Burton & Schwarz, 2013) or payments by modeled results (Bartkowski et al., 2021).

What is more, Studies 1, 2 and 4 highlighted the importance of technology and (eco-)technical change. Legislators could address this particular point to foster productive, as well as environmentally and climate-friendly farming practices. Beltrán-Esteve & Picazo-Tadeo (2015) suggest that the stimulation of eco-innovations could improve the ecological-economic performance of businesses. Long et al. (2016) recommend, among other things, financial support for start-up companies and tax-cuts for research and development activities. This could boost technological improvements and have a positive impact on farms' emission-performance and finally on their relative climate change mitigation potential. In terms of agroforestry adoption, reducing the current minimum useful lifetime of the wood-based land use options could increase adoption. Legislators could establish a framework to encourage the development of coppices with reduced minimum useful lifetime but without reduced economic benefits. One way to do this might be the promotion of novel breeding methods, which have shown high innovation potential across several domains (Qaim, 2020).

Another approach is to internalize environmental externalities induced by farming activities. By more effectively conditioning farmers' income to their climate-protection and environmental performance, a behavior which is more oriented towards the public good can be expected (compare Beltrán-Esteve et al., 2014; Picazo-Tadeo et al., 2012). E.g. Picazo-Tadeo et al. (2011) demand a stronger commitment of EU policy-makers to the principle of *conditionality*, i.e. only farmers that comply with ambitious ecological standards should benefit from public resources. Following this line of argumention, European legislators have introduced the above-mentioned eco-schemes, a novel set of policy instruments that is supposed to strengthen the conditionality principle (Latacz-Lohmann et al., 2022). If this instrument will be effective in this regard remains to be seen and depends on the actual implementation of the EU member states.

This thesis has shown that it is possible to reconcile environmental goals (specifically climate change mitigation) and economic returns. In light of this, current claims to increase agricultural production at the expense of environmental objectives due to the Russian invasion in Ukraine and the associated rise in agricultural commodity prices (Baffes & Nagle, 2022) should be looked at critically. For instance, the European Commission has allowed farmers to grow food on fallow land, which has originally been dedicated to increase biodiversity (Blenkinsop & Baczynska, 2022). Giving up environmental protection in favor of a (marginal) production boost is questionable. Studies 1 and 2 of this thesis have shown that agricultural production and environmental protection are not necessarily mutually exclusive. Legislators should therefore try to promote the eradication of eco-inefficiencies before sacrificing environmental protection areas in favor of production.

8.7 Future research directions

The analysis of the relationship between agricultural production and environmental and climate change is an active research field. There are interesting pathways for future research resulting from this thesis.

First, it would be interesting to analyze more environmental pressures and longer time horizons in the realm of the eco-efficiency studies. This would allow a more holistic assessment of the ecological-economic performance of farms. It would also be interesting to conduct a spatially explicit eco-performance analysis to detect local or regional (in-)efficiency hotspots. In the presence of spatial information, it would also be possible to account for spatial correlation through a spatial autoregressive eco-efficiency model.

Second, regarding the criticisms on the eco-efficiency approach from an axiomatic perspective, future research could focus on methodological advances regarding the combination of axiomatically more consistent by-production models (Murty et al., 2012) and latent-class analysis (Orea & Kumbhakar, 2004) in the spirit of Study 2, which would further improve the understanding of the production-environment nexus.

Third, more work is needed to elaborate on the usefulness of different causal machine learning methods and the inclusion of novel data sources (e.g. remotely sensed) in the context of agricultural impact analysis to further explore their strengths and weaknesses in comparison to more traditional approaches. This is also true when it comes to exploring the possibilities of using model-agnostic interpretation methods in the realm of causal machine learning approaches.

Fourth, an important extension to the analysis in Study 3 would be the assessment of subprogram-specific heterogeneous treatment effects. If there was information on specific agri-environmental subprograms, it might be possible to look at specific schemes individually. Heiler & Knaus (2021) propose a promising nonparametric decomposition method for the estimation and statistical inference of effect heterogeneity and treatment heterogeneity.

Fifth, we cannot observe the effect of AES over time as we are restricted to one year in our analysis. As farms, however, must generally participate for a period of at least five years, we might miss important temporal structures as well as lagged and build-up effects of agri-environment measures. An extension of the method presented in Study 3 to account for these effects would be a promising direction for future research. Miller (2020) offers a good starting point for such an analysis.

Finally, to better understand and compare farmers' responses to extreme weather events, it would be interesting to see similar analyses to the one in Study 4 assessing other environmentally and climate-friendly land uses and technologies.

References

Addo, F. & Salhofer, K. (2022). Transient and persistent technical efficiency and its determinants: the case of crop farms in Austria. *Applied Economics*, 54(25), 2916–2932.

Adesina, A. A. & Chianu, J. (2002). Determinants of farmers' adoption and adaptation of alley farming technology in Nigeria. *Agroforestry Systems*, 55(2), 99–112.

Ahn, S. C. & Sickles, R. C. (2000). Estimation of long-run inefficiency levels: a dynamic frontier approach. *Econometric Reviews*, 19(4), 461–492.

Aigner, D., Lovell, C., & Schmidt, P. (1977). Formulation and estimation of stochastic frontier production function models. *Journal of Econometrics*, 6(1), 21–37.

Ait Sidhoum, A., Dakpo, K. H., & Latruffe, L. (2022). Trade-offs between economic, environmental and social sustainability on farms using a latent class frontier efficiency model: Evidence for Spanish crop farms. *PLOS ONE*, 17(1), 1–17.

Ait Sidhoum, A., Serra, T., & Latruffe, L. (2019). Measuring sustainability efficiency at farm level: a data envelopment analysis approach. *European Review of Agricultural Economics*, 47(1), 1–26.

Albrecht, A. & Kandji, S. T. (2003). Carbon sequestration in tropical agroforestry systems. *Agriculture, Ecosystems and Environment*, 99(1-3), 15–27.

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.

Alvarez, A., Del Corral, J., & Tauer, L. W. (2012). Modeling unobserved heterogeneity in New York dairy farms: One-stage versus two-stage models. *Agricultural and Resource Economics Review*, 41(3), 275–285.

Amare, D. & Darr, D. (2020). Agroforestry adoption as a systems concept: A review. *Forest Policy and Economics*, 120(July), 102299.

Amusa, T. & Simonyan, J. (2018). Discrete Choice Model of Farmers' Willingness-to Engage in Agroforestry-Based Farming in Southwest, Nigeria. *Earth and Atmospheric Sciences, American Association for Science and Technology*, 1(2), 46–54.

Andersen, S., Harrison, G. W., Lau, M. I., & Elisabet Rutström, E. (2008). Lost in state space: Are preferences stable? *International Economic Review*, 49(3), 1091–1112.

Angulo-Meza, L., González-Araya, M., Iriarte, A., Rebolledo-Leiva, R., & Soares de Mello, J. C. (2019). A multiobjective DEA model to assess the eco-efficiency of agricultural practices within the CF+DEA method. *Computers and Electronics in Agriculture*, 161(October 2017), 151–161.

Arata, L. & Sckokai, P. (2016). The impact of agri-environmental schemes on farm performance in five E.U. member States: A DID-matching approach. *Land Economics*, 92(1), 167–186.

Arbia, G. (1989). *Spatial Data Configuration in Statistical Analysis of Regional Economic and Related Problems*, volume 14 of *Advanced Studies in Theoretical and Applied Econometrics*. Dordrecht: Springer Netherlands.

Armsworth, P. R., Acs, S., Dallimer, M., Gaston, K. J., Hanley, N., & Wilson, P. (2012). The cost of policy simplification in conservation incentive programs. *Ecology letters*, 15(5), 406–414.

ART (2016). *Ex post-Bewertung des Bayerischen Zukunftsprogramms Agrarwirtschaft und Ländlicher Raum 2007-2013 (BayZAL)*. Weidenbach-Triesdorf: Forschungsgruppe Agrar- und Regionalentwicklung Triesdorf (ART).

ART (2019). *Bewertung des Entwicklungsprogramms für den ländlichen Raum in Bayern 2014 – 2020 (EPLR Bayern 2020) Maßnahmenspezifische Bewertung*. Weidenbach-Triesdorf: Forschungsgruppe Agrar- und Regionalentwicklung Triesdorf (ART).

Ashraf, M., Sanusi, R., Zulkifli, R., Tohiran, K. A., Moslim, R., Ashton-Butt, A., & Azhar, B. (2019). Alley-cropping system increases vegetation heterogeneity and moderates extreme microclimates in oil palm plantations. *Agricultural and Forest Meteorology*, 276-277(July), 107632.

Athey, S. & Imbens, G. (2016). Recursive partitioning for heterogeneous causal effects. *Proceedings of the National Academy of Sciences of the United States of America*, 113(27), 7353–7360.

Athey, S., Imbens, G., Pham, T., & Wager, S. (2017). Estimating average treatment effects: Supplementary analyses and remaining challenges. *American Economic Review*, 107(5), 278–281.

Athey, S. & Imbens, G. W. (2019). Machine Learning Methods That Economists Should Know About. *Annual Review of Economics*, 11(1), 685–725.

Athey, S., Tibshirani, J., & Wager, S. (2019). Generalized random forests. *Annals of Statistics*, 47(2), 1179–1203.

Athey, S. & Wager, S. (2019). Estimating Treatment Effects with Causal Forests: An Application. URL: http://arxiv.org/abs/1902.07409 (last accessed: 2021-12-01).

Atkinson, S. E., Primont, D., & Tsionas, M. G. (2018). Statistical inference in efficient production with bad inputs and outputs using latent prices and optimal directions. *Journal of Econometrics*, 204(2), 131–146.

Atkinson, S. E. & Tsionas, M. G. (2018). Shadow directional distance functions with bads: GMM estimation of optimal directions and efficiencies. *Empirical Economics*, 54(1), 207–230.

Austin, E. J., Deary, I. J., & Willock, J. (2001). Personality and intelligence as predictors of economic behaviour in Scottish farmers. *European Journal of Personality*, 15, S123–S137.

Avohou, T. H., Houehounha, R., Glele-Kakai, R., Assogbadjo, A. E., & Sinsin, B. (2011). Firewood yield and profitability of a traditional Daniellia oliveri short-rotation coppice on fallow lands in Benin. *Biomass and Bioenergy*, 35(1), 562–571.

Baer, S. G., Engle, D. M., Knops, J. M. H., Langeland, K. A., Maxwell, B. D., Menalled, F. D., & Symstad, A. J. (2009). Vulnerability of rehabilitated agricultural production systems to invasion by nontarget plant species. *Environmental management*, 43(2), 189–196.

Baffes, J. & Nagle, P. (2022). Commodity prices surge due to the war in Ukraine. In *World Bank Blogs*. World Bank. URL: blogs.worldbank.org/developmenttalk/commodity-prices-surge-due-war-ukraine (last accessed: 2022-27-05).

Baldoni, E., Coderoni, S., & Esposti, R. (2017). The productivity and environment nexus with farm-level data. The case of carbon footprint in lombardy FADN farms. *Bio-based and Applied Economics*, 6(2), 119–137.

Baldoni, E., Coderoni, S., & Esposti, R. (2018). The complex farm-level relationship between environmental performance and productivity : The case of carbon footprint of Lombardy farms. *Environmental Science and Policy*, 89(December 2017), 73–82.

Balk, B. M. (2001). Scale Efficiency and Productivity Change. *Journal of Productivity Analysis*, 15(October 1999), 159–183.

Ball, V. E., Lovell, C. A. K., Luu, H., & Nehring, R. (2004). Environmental Incorporating of Impacts in the Measurement Growth Productivity Agricultural. *Journal of Agricultural and Resource Economic*, 29(3), 436–460.

Bannister, M. E. & Nair, P. K. (2003). Agroforestry adoption in Haiti: The importance of household and farm characteristics. *Agroforestry Systems*, 57(2), 149–157.

Baráth, L., Fertő, I., & Bojnec, Š. (2020). The Effect of Investment, LFA and Agrienvironmental Subsidies on the Components of Total Factor Productivity: The Case of Slovenian Farms. *Journal of Agricultural Economics*, 71(3), 853–876.

Barghusen, R., Sattler, C., Deijl, L., Weebers, C., & Matzdorf, B. (2021). Motivations of farmers to participate in collective agri-environmental schemes: the case of Dutch agricultural collectives. *Ecosystems and People*, 17(1), 539–555.

Bartkowski, B., Droste, N., Ließ, M., Sidemo-Holm, W., Weller, U., & Brady, M. V. (2021). Payments by modelled results: A novel design for agri-environmental schemes. *Land Use Policy*, 102(November 2020), 105230.

Bartolini, F., Vergamini, D., Longhitano, D., & Povellato, A. (2021). Do differential payments for agri-environment schemes affect the environmental benefits? A case study in the North-Eastern Italy. *Land Use Policy*, 107(April 2019), 104862.

Basset-Mens, C., Ledgard, S., & Boyes, M. (2009). Eco-efficiency of intensification scenarios for milk production in New Zealand. *Ecological Economics*, 68(6), 1615–1625.

Batáry, P., Dicks, L. V., Kleijn, D., & Sutherland, W. J. (2015). The role of agri-environment schemes in conservation and environmental management. *Conservation biology : the journal of the Society for Conservation Biology*, 29(4), 1006–1016.

Battese, G. E. & Coelli, T. J. (1988). Prediction of firm-level technical efficiencies with a generalized frontier production function and panel data. *Journal of Econometrics*, 38(3), 387–399.

Battese, G. E. & Coelli, T. J. (1992). Frontier production functions, technical efficiency and panel data: With application to paddy farmers in India. *Journal of Productivity Analysis*, 3(1-2), 153–169.

Battese, G. E. & Coelli, T. J. (1995). A model for technical inefficiency effects in a stochastic frontier production function for panel data. *Empirical Economics*, 20(2), 325–332.

Battocchi, K., Dillon, E., Hei, M., Lewis, G., Oka, P., Oprescu, M., & Syrgkanis, V. (2019). EconML: A Python Package for ML-Based Heterogeneous Treatment Effects Estimation. URL: github.com/microsoft/EconML (last accessed: 2021-5-22).

Baum, S., Weih, M., Busch, G., Koiher, F., & Bolte, A. (2009). The impact of Short Rotation Coppice plantations on phytodiversity. *Landbauforschung - vTI Agriculture and Forestry Research*, 59(3), 159–162.

Bayard, B., Jolly, C. M., & Shannon, D. A. (2007). The economics of adoption and management of alley cropping in Haiti. *Journal of Environmental Management*, 84(1), 62–70.

Bayerisches Landesamt für Statistik (2022). Land- und Forstwirtschaft, Fischerei. Bavarian State Agency for Statistics and Data Processing. URL: https://www.statistikdaten.bayern.de/genesis/online?operation=themes&code=4#abreadcrumb (last accessed: 2022-07-06).

Bellebaum, J. & Koffijberg, K. (2018). Present agri-environment measures in Europe are not sufficient for the conservation of a highly sensitive bird species, the Corncrake Crex crex. *Agriculture, Ecosystems & Environment*, 257, 30–37.

Beltrán-Esteve, M., Gómez-Limón, J. A., Picazo-Tadeo, A. J., & Reig-Martínez, E. (2014). A metafrontier directional distance function approach to assessing eco-efficiency. *Journal of Productivity Analysis*, 41(1), 69–83.

Beltrán-Esteve, M. & Picazo-Tadeo, A. J. (2015). Assessing environmental performance trends in the transport industry: Eco-innovation or catching-up? *Energy Economics*, 51, 570–580.

Beltrán-Esteve, M. & Picazo-Tadeo, A. J. (2017). Assessing environmental performance in the European Union: Eco-innovation versus catching-up. *Energy Policy*, 104(January), 240–252.

Béné, C., Wood, R. G., Newsham, A., & Davies, M. (2012). Resilience: New Utopia or New Tyranny? Reflection about the Potentials and Limits of the Concept of Resilience in Relation to Vulnerability Reduction Programmes. *IDS Working Papers*, 2012(405), 1–61.

Bennett, A. & Kallus, N. (2019). Policy evaluation with latent confounders via optimal balance. *33rd Conference on Neural Information Processing Systems (NeurIPS 2019)*, URL: http:.

Benton, T. G., Vickery, J. A., & Wilson, J. D. (2003). Farmland biodiversity: Is habitat heterogeneity the key? *Trends in Ecology and Evolution*, 18(4), 182–188.

Bertoni, D., Aletti, G., Cavicchioli, D., Micheletti, A., & Pretolani, R. (2021). Estimating the CAP greening effect by machine learning techniques: A big data ex post analysis. *Environmental Science and Policy*, 119, 44–53.

Bertoni, D., Curzi, D., Aletti, G., & Olper, A. (2020). Estimating the effects of agri-environmental measures using difference-in-difference coarsened exact matching. *Food Policy*, 90(November), 101790.

Beyene, A. D., Mekonnen, A., Randall, B., & Deribe, R. (2019). Household Level Determinants of Agroforestry Practices Adoption in Rural Ethiopia. *Forests Trees and Livelihoods*, 28(3), 194–213.

BGR (2022). *Karte der Bodenarten in Oberböden 1:1.000.000 (BOART1000OB).* URL: www.bgr.bund.de/DE/Themen/Boden/Informationsgrundlagen/Bodenkundliche_-Karten_Datenbanken/Themenkarten/BOART1000OB/boart1000ob_-node.html (last accessed: 2022-4-27): Bundesanstalt fuer Geowissenschaften und Rohstoffe.

Biermann, U., Bornscheuer, U., Meier, M. A., Metzger, J. O., & Schäfer, H. J. (2011). Oils and fats as renewable raw materials in chemistry. *Angewandte Chemie - International Edition*, 50(17), 3854–3871.

Billor, N., Hadi, A. S., & Velleman, P. F. (2000). BACON: Blocked adaptive computationally efficient outlier nominators. *Computational Statistics and Data Analysis*, 34(3), 279–298.

Birge, T., Toivonen, M., Kaljonen, M., & Herzon, I. (2017). Probing the grounds: Developing a payment-by-results agri-environment scheme in Finland. *Land Use Policy*, 61, 302–315.

Blenkinsop, P. & Baczynska, G. (2022). EU offers farmers aid, more land to grow due to Ukraine war. URL: www.reuters.com/world/eu-offers-farmers-aid-more-land-grow-due-ukraine-war-2022-03-23/ (last accessed: 2022-27-05): Reuters.

Bogetoft, P. & Hougaard, J. L. (2003). Rational Inefficiencies. *Journal of Productivity Analysis*, 20(3), 243–271.

Bonfiglio, A., Arzeni, A., & Bodini, A. (2017). Assessing eco-efficiency of arable farms in rural areas. *Agricultural Systems*, 151, 114–125.

Borges, J. A. R., Foletto, L., & Xavier, V. T. (2015). An interdisciplinary framework to study farmers ' decisions on adoption of innovation : Insights from Expected Utility Theory and Theory of Planned Behavior. 10(29), 2814–2825.

Bougherara, D., Lapierre, M., Préget, R., & Sauquet, A. (2021). Do farmers prefer increasing, decreasing, or stable payments in Agri-environmental schemes? *Ecological Economics*, 183(November 2020), 106946.

Breiman, L. (2001). Random Forests. *Machine Learning*, 45(5-32), 1–122.

Breiman, L., Friedman, J. H., Olshen, R. A., & Stone, C. J. (2017). *Classification And Regression Trees*. The Wadsworth and Brooks-Cole statistics-probability series. New York: Routledge.

Briec, W. & Kerstens, K. (2011). The hicks-moorsteen productivity index satisfies the determinateness axiom. *Manchester School*, 79(4), 765–775.

Bright, J. A., Morris, A. J., Field, R. H., Cooke, A. I., Grice, P. V., Walker, L. K., Fern, J., & Peach, W. J. (2015). Higher-tier agri-environment scheme enhances breeding densities of some priority farmland birds in England. *Agriculture, Ecosystems & Environment*, 203, 69–79.

Brown, S. E., Miller, D. C., Ordonez, P. J., & Baylis, K. (2018). Evidence for the impacts of agroforestry on agricultural productivity, ecosystem services, and human well-being in high-income countries: A systematic map protocol. *Environmental Evidence*, 7(1), 1–16.

Brussaard, L., de Ruiter, P. C., & Brown, G. G. (2007). Soil biodiversity for agricultural sustainability. *Agriculture, Ecosystems and Environment*, 121(3), 233–244.

Burke, M. & Emerick, K. (2015). Adaptation to climate change: Evidence from US agriculture. *American Economic Journal: Economic Policy*, 8(3), 106–140.

Burke, M., Hsiang, S. M., & Miguel, E. (2015). Global non-linear effect of temperature on economic production. *Nature*, 527(7577), 235–239.

Burney, J. A., Davis, S. J., & Lobell, D. B. (2010). Greenhouse gas mitigation by agricultural intensification. *Proceedings of the National Academy of Sciences of the United States of America*, 107(26), 12052–12057.

Burton, R. J. & Schwarz, G. (2013). Result-oriented agri-environmental schemes in Europe and their potential for promoting behavioural change. *Land Use Policy*, 30(1), 628–641.

Buyinza, J., Nuberg, I. K., Muthuri, C. W., & Denton, M. D. (2020). Assessing smallholder farmers' motivation to adopt agroforestry using a multi-group structural equation modeling approach. *Agroforestry Systems*, 94(6), 2199–2211.

Calvi, G., Campedelli, T., Tellini Florenzano, G., & Rossi, P. (2018). Evaluating the benefits of agri-environment schemes on farmland bird communities through a common species monitoring programme. A case study in northern Italy. *Agricultural Systems*, 160, 60–69.

Camarero, M., Castillo-Giménez, J., Picazo-Tadeo, A. J., & Tamarit, C. (2014). Is eco-efficiency in greenhouse gas emissions converging among European Union countries? *Empirical Economics*, 47(1), 143–168.

Campbell, B. M., Beare, D. J., Bennett, E. M., Hall-Spencer, J. M., Ingram, J. S. I., Jaramillo, F., Ortiz, R., Ramankutty, N., Sayer, J. A., & Shindell, D. (2017). Agriculture production as a major driver of the Earth system exceeding planetary boundaries. *Ecology and Society*, 22(4), Art. 8.

Capper, J. L., Cady, R. A., & Bauman, D. E. (2009). The environmental impact of dairy production: 1944 compared with 20071. *Journal of Animal Science*, 87(6), 2160–2167.

Cardinael, R., Cadisch, G., Gosme, M., Oelbermann, M., & van Noordwijk, M. (2021). Climate change mitigation and adaptation in agriculture: Why agroforestry should be part of the solution. *Agriculture, Ecosystems & Environment*, 319(June), 107555.

Cardinael, R., Chevallier, T., Cambou, A., Béral, C., Barthès, B. G., Dupraz, C., Durand, C., Kouakoua, E., & Chenu, C. (2017). Increased soil organic carbon stocks under agroforestry: A survey of six different sites in France. *Agriculture, Ecosystems and Environment*, 236, 243–255.

Carter, M. R., Tjernström, E., & Toledo, P. (2019). Heterogeneous impact dynamics of a rural business development program in Nicaragua. *Journal of Development Economics*, 138, 77–98.

Casey, J. W. & Holden, N. M. (2005). Analysis of greenhouse gas emissions from the average Irish milk production system. *Agricultural Systems*, 86(1), 97–114.

Caves, D. W., Christensen, L. R., & Diewert, W. E. (1982). The Economic Theory of Index Numbers and the Measurement of Input, Output, and Productivity. *Econometrica*, 50(6), 1393–1414.

Chabé-Ferret, S. & Subervie, J. (2013). How much green for the buck? Estimating additional and windfall effects of French agro-environmental schemes by DID-matching. *Journal of Environmental Economics and Management*, 65(1), 12–27.

Chambers, R. G. (1988). *Applied Production Analysis*. New York: Cambridge University Press.

Chambers, R. G. & Lee, H. (1986). Constrained output maximization and US agriculture. *Applied Economics*, 18(4), 347–357.

Chambers, R. G. & Pieralli, S. (2020). The Sources of Measured US Agricultural Productivity Growth: Weather, Technological Change, and Adaptation. *American Journal of Agricultural Economics*, 102(4), 1198–1226.

Chambers, R. G., Pieralli, S., & Sheng, Y. (2020). The Millennium Droughts and Australian Agricultural Productivity Performance: A Nonparametric Analysis. *American Journal of Agricultural Economics*, 102(5), 1383–1403.

Chen, Y. Y., Schmidt, P., & Wang, H. J. (2014). Consistent estimation of the fixed effects stochastic frontier model. *Journal of Econometrics*, 181(2), 65–76.

Chernozhukov, V., Chetverikov, D., Demirer, M., Duflo, E., Hansen, C., Newey, W., & Robins, J. (2018). Double/debiased machine learning for treatment and structural parameters. *Econometrics Journal*, 21(1), C1–C68.

Chi, M., Plaza, A., Benediktsson, J. A., Sun, Z., Shen, J., & Zhu, Y. (2016). Big Data for Remote Sensing: Challenges and Opportunities. *Proceedings of the IEEE*, 104(11), 2207–2219.

Ciais, P., Reichstein, M., Viovy, N., Granier, A., Ogée, J., Allard, V., Aubinet, M., Buchmann, N., Bernhofer, C., Carrara, A., Chevallier, F., De Noblet, N., Friend, A. D., Friedlingstein, P., Grünwald, T., Heinesch, B., Keronen, P., Knohl, A., Krinner, G., Loustau, D., Manca, G., Matteucci, G., Miglietta, F., Ourcival, J. M., Papale, D., Pilegaard, K., Rambal, S., Seufert, G., Soussana, J. F., Sanz, M. J., Schulze, E. D., Vesala, T., & Valentini, R. (2005). Europe-wide reduction in primary productivity caused by the heat and drought in 2003. *Nature*, 437(7058), 529–533.

Cinelli, C., Forney, A., & Pearl, J. (2020). A Crash Course in Good and Bad Controls. *SSRN*, (URL: www.ssrn.com/abstract=3689437 (last accessed: 2022-2-16)).

Coderoni, S. & Esposti, R. (2014). Is There a Long-Term Relationship Between Agricultural GHG Emissions and Productivity Growth? A Dynamic Panel Data Approach. *Environmental and Resource Economics*, 58(2), 273–302.

Coderoni, S. & Esposti, R. (2018). CAP payments and agricultural GHG emissions in Italy. A farm-level assessment. *Science of the Total Environment*, 627, 427–433.

Coelli, T. J. & Rao, D. S. (2005a). Total factor productivity growth in agriculture: A Malmquist index analysis of 93 countries, 1980-2000. *Agricultural Economics*, 32(SUPPL. 1), 115–134.

Coelli, T. J. & Rao, D. S. P. (2005b). Total factor productivity growth in agriculture: a Malmquist index analysis of 93 countries, 1980–2000. *Agricultural Economics*, 32(s1), 115–134.

Coluccia, B., Valente, D., Fusco, G., De Leo, F., & Porrini, D. (2020). Assessing agricultural eco-efficiency in Italian Regions. *Ecological Indicators*, 116(March), 106483.

Cornes, R. C., van der Schrier, G., van den Besselaar, E. J. M., & Jones, P. D. (2018). An Ensemble Version of the E-OBS Temperature and Precipitation Data Sets. *Journal of Geophysical Research: Atmospheres*, 123(17), 9391–9409.

Cortés, A., Feijoo, G., Fernández, M., & Moreira, M. T. (2021). Pursuing the route to eco-efficiency in dairy production: The case of Galician area. *Journal of Cleaner Production*, 285, 124861.

Crosson, P., Shalloo, L., O'Brien, D., Lanigan, G. J., Foley, P. A., Boland, T. M., & Kenny, D. A. (2011). A review of whole farm systems models of greenhouse gas emissions from beef and dairy cattle production systems. *Animal Feed Science and Technology*, 166-167, 29–45.

Cui, X. (2020). Climate change and adaptation in agriculture: Evidence from US cropping patterns. *Journal of Environmental Economics and Management*, 101, 102306.

Cui, X., Zhou, F., Ciais, P., Davidson, E. A., Tubiello, F. N., Niu, X., Ju, X., Canadell, J. G., Bouwman, A. F., Jackson, R. B., Mueller, N. D., Zheng, X., Kanter, D. R., Tian, H., Adalibieke, W., Bo, Y., Wang, Q., Zhan, X., & Zhu, D. (2021). Global mapping of crop-specific emission factors highlights hotspots of nitrous oxide mitigation. *Nature Food*, 2(11), 886–893.

Cunningham, S. (2021). *Causal Inference: The Mixtape*. Yale University Press.

Cuong, N. V. (2009). Impact evaluation of multiple overlapping programs under a conditional independence assumption. *Research in Economics*, 63(1), 27–54.

Dadam, D. & Siriwardena, G. M. (2019). Agri-environment effects on birds in Wales: Tir Gofal benefited woodland and hedgerow species. *Agriculture, Ecosystems & Environment*, 284, 106587.

Dakpo, K. H., Jeanneaux, P., & Latruffe, L. (2017). Greenhouse gas emissions and efficiency in French sheep meat farming: A nonparametric framework of pollutionadjusted technologies. *European Review of Agricultural Economics*, 44(1), 33–65.

Dakpo, K. H., Jeanneaux, P., & Latruffe, L. (2020). Modelling Pollution-Generating Technologies: A Numerical Comparison of Non-parametric Approaches. In J. Aparicio, C. A. Knox Lovell, J. T. Pastor, & J. Zhu (Eds.), *Advances in Efficiency and Productivity II* (pp. 67–85). Cambridge: Springer.

Dakpo, K. H. & Lansink, A. O. (2019). Dynamic pollution-adjusted inefficiency under the by-production of bad outputs. *European Journal of Operational Research*, 276(1), 202–211.

Dal Ferro, N., Cocco, E., Lazzaro, B., Berti, A., & Morari, F. (2016). Assessing the role of agri-environmental measures to enhance the environment in the Veneto Region, Italy, with a model-based approach. *Agriculture, Ecosystems & Environment*, 232, 312–325.

Dalhaus, T., Schlenker, W., Blanke, M. M., Bravin, E., & Finger, R. (2020). The Effects of Extreme Weather on Apple Quality. *Scientific Reports*, 10(1), 1–7.

Daly, A., Hess, S., & Train, K. (2012). Assuring finite moments for willingness to pay in random coefficient models. *Transportation*, 39(1), 19–31.

Dasgupta, U., Gangadharan, L., Maitra, P., & Mani, S. (2017). Searching for preference stability in a state dependent world. *Journal of Economic Psychology*, 62, 17–32.

De Giusti, G., Kristjanson, P., & Rufino, M. C. (2019). Agroforestry as a climate change mitigation practice in smallholder farming: evidence from Kenya. *Climatic Change*, 153(3), 379–394.

Debertin, D. L. (2012). *Agricultural Production Economics*. Lexington, second edition.

Defrancesco, E., Gatto, P., & Mozzato, D. (2018). To leave or not to leave? Understanding determinants of farmers' choices to remain in or abandon agri-environmental schemes. *Land Use Policy*, 76(August 2017), 460–470.

Deines, J. M., Wang, S., & Lobell, D. B. (2019). Satellites reveal a small positive yield effect from conservation tillage across the US Corn Belt. *Environmental Research Letters*, 14(12), 124038.

Del Rossi, G., Hecht, J. S., & Zia, A. (2021). A mixed-methods analysis for improving farmer participation in agri-environmental payments for ecosystem services in Vermont, USA. *Ecosystem Services*, 47(May 2019), 101223.

Dell, M., Jones, B. F., & Olken, B. A. (2014). What Do We Learn from the Weather? The New Climate-Economy Literature. *Journal of Economic Literature*, 52(3), 740–798.

den Herder, M., Moreno, G., Rosa Mosquera-Losada, Palma, J., Sidiropoulou, A., Santiago Freijanes, J. J., Crous-Duran, J., Paulo, J., Tomé, M., Pantera, A., Papanastasis, V., Mantzanas, K., Pachana, P., & Burgess, P. J. (2015). Current extent and trends of agroforestry in the EU27. URL: www.agforward.eu/index.php/en/current-extent-.

Deschênes, O. & Greenstone, M. (2007). The economic impacts of climate change: Evidence from agricultural output and random fluctuations in weather. *American Economic Review*, 97(1), 354–385.

Deschênes, O. & Greenstone, M. (2012). The economic impacts of climate change: Evidence from agricultural output and random fluctuations in weather: Reply. *American Economic Review*, 102(7), 3761–3773.

Desjeux, Y., Dupraz, P., Kuhlman, T., Paracchini, M. L., Michels, R., Maigné, E., & Reinhard, S. (2015). Evaluating the impact of rural development measures on nature value indicators at different spatial levels: Application to France and The Netherlands. *Ecological Indicators*, 59, 41–61.

Dessart, F. J., Barreiro-Hurlé, J., & Van Bavel, R. (2019). Behavioural factors affecting the adoption of sustainable farming practices: A policy-oriented review. *European Review of Agricultural Economics*, 46(3), 417–471.

Destatis (2021a). Förderprogramme Landwirtschaftszählung 2020. URL: https://www.destatis.de/DE/Themen/Branchen-Unternehmen/Landwirtschaft-Forstwirtschaft-Fischerei/Landwirtschaftliche-Betriebe/Publikationen/Downloads-Landwirtschaftliche-Betriebe/foerderprogramme-5411206209004.html (last accessed: 2022-06-07).

Destatis (2021b). Landwirtschaftszählung 2020 - Regionale Ergebnisse. URL: https://www.statistikportal.de/de/veroeffentlichungen/landwirtschaftszaehlung-2020 (last accessed: 2022-06-07).

Deutscher Bundestag (2021). Beschlussempfehlung und Bericht des Ausschusses für Ernährung und Landwirtschaft (10. Ausschuss) Drucksache 19/30513 (neu). URL: dserver.bundestag.de/btd/19/305/1930513.pdf (last accessed: 2022-5-26).

Dhakal, A., Cockfield, G., & Maraseni, T. N. (2015). Deriving an index of adoption rate and assessing factors affecting adoption of an agroforestry-based farming system in Dhanusha District, Nepal. *Agroforestry Systems*, 89(4), 645–661.

Di Falco, S., Adinolfi, F., Bozzola, M., & Capitanio, F. (2014). Crop Insurance as a Strategy for Adapting to Climate Change. *Journal of Agricultural Economics*, 65(2), 485–504.

Dick, J., Smith, P., Smith, R., Lilly, A., Moxey, A., Booth, J., Campbell, C., & Coulter, D. (2008). Calculating farm scale greenhouse gas emissions. (June), 0–29.

Dillon, J. L. (1992). The Farm as a Purposeful System. In *Department of Agricultural Economics & Business Management, Miscellaneous Publication No. 10.* Armidale: The University of New England.

DiPrete, T. A. & Gangl, M. (2004). Assessing bias in the estimation of causal effects: Rosenbaum bounds on matching estimators and instrumental variables estimation with imperfect instruments. *Sociological Methodology*, 34(1), 271–310.

Djalilov, B. M., Khamzina, A., Hornidge, A. K., & Lamers, J. P. (2016). Exploring constraints and incentives for the adoption of agroforestry practices on degraded cropland in Uzbekistan. *Journal of Environmental Planning and Management*, 59(1), 142–162.

Doiron, D. & Yoo, H. I. (2017). Temporal Stability of Stated Preferences: The Case of Junior Nursing Jobs. *Health Economics*, 26(6), 802–809.

Doove, L. L., Van Buuren, S., & Dusseldorp, E. (2014). Recursive partitioning for missing data imputation in the presence of interaction effects. *Computational Statistics and Data Analysis*, 72, 92–104.

Duguma, L. A., Minang, P. A., & Van Noordwijk, M. (2014). Climate change mitigation and adaptation in the land use sector: From complementarity to synergy. *Environmental Management*, 54(3), 420–432.

Dumangane, M., Freo, M., Granato, S., Lapatinas, A., & Mazzarella, G. (2021). *An evaluation of the CAP impact: A Discrete Policy Mix Analysis*. Luxembourg: Publications Office of the European Union.

Dundas, S. J. & von Haefen, R. H. (2020). The Effects of Weather on Recreational Fishing Demand and Adaptation: Implications for a Changing Climate. *Journal of the Association of Environmental and Resource Economists*, 7(2), 209–242.

Dupraz, P. & Guyomard, H. (2019). Environment and Climate in the Common Agricultural Policy. *EuroChoices*, 18(1), 18–25.

DWD (2022). *German Climate Atlas Explanations*. URL: www.dwd.de/EN/ourservices/germanclimateatlas/explanations/elements/_-functions/faqkarussel/heissetage.html (last accessed: 2021-12-12): Germany's National Meteorological Service -Deutscher Wetterdienst.

Eder, A. (2022). Environmental efficiency measurement when producers control pollutants under heterogeneous conditions: a generalization of the materials balance approach. *Journal of Productivity Analysis*, 57(2), 157–176.

Eder, A., Salhofer, K., & Scheichel, E. (2021). Land tenure, soil conservation, and farm performance: An eco-efficiency analysis of Austrian crop farms. *Ecological Economics*, 180(August 2020), 106861.

El-Nazer, T. & McCarl, B. A. (1986). The Choice of Crop Rotation: A Modeling Approach and Case Study. *American Journal of Agricultural Economics*, 68(1), 127–136.

Emvalomatis, G., Stefanou, S. E., & Lansink, A. O. (2011). A Reduced-Form Model for Dynamic Efficiency Measurement: Application to Dairy Farms in Germany and The Netherlands. *American Journal of Agricultural Economics*, 93(1), 161–174.

Engle, N. L. (2011). Adaptive capacity and its assessment. *Global Environmental Change*, 21(2), 647–656.

Espinosa-Goded, M., Barreiro-Hurlé, J., & Dupraz, P. (2013). Identifying additional barriers in the adoption of agri-environmental schemes: The role of fixed costs. *Land Use Policy*, 31, 526–535.

ETCCDI (2018). *Climate Change Indices*. URL: http://etccdi.pacificclimate.org/list_27_indices.shtml (last accessed: 2021-08-05): Expert Team on Climate Change Detection and Indices.

European Commission (2013). Regulation (EU) No 1305/2013 of the European Parliament and of the Council of 17 december 2013 on support for rural development by the European Agricultural Fund for Rural Development (EAFRD) and repealing Council Regulation (EC) No 1698/2005. Brussels: European Commission. Brussels.

European Commission (2018a). Proposal for a regulation of the European Parliament and Council amending Regulations (EU) No 1308/2013 establishing a common organisation of the markets in agricultural products. Brussels: European Commission.

European Commission (2018b). Proposal for a regulation of the European Parliament and Council establishing rules on support for strategic plans to be drawn up by Member States under the Common agricultural policy (CAP Strategic Plans). Brussels: European Commission.

European Commission (2018c). Proposal for a regulation of the European Parliament and the Council on the financing, management and monitoring of the common agricultural policy and repealing Regulation (EU) No 1306/2013. Brussels: European Commission.

European Commission (2019a). *CAP Specific Objectives … explained – Brief No 4*. Technical Report 1, Brussels.

European Commission (2019b). *COP21: United for Climate*. Agricultural Policy. Brussels: European Commission.

European Commission (2019c). *Increasing Competitiveness: the Role of Productivity*. CAP specific objectives ...explained - Brief No 2. Brussels: European Commission.

European Commission (2019d). The Post-2020 Common Agricultural Policy: Environmental Benefits and Simplification. Brussels.

European Commission (2022a). Conversion to a Farm Sustainability Data Network (FSDN). URL: ec.europa.eu/info/law/better-regulation/have-your-say/initiatives/12951-Conversion-to-a-Farm-Sustainability-Data-Network-FSDN-_en (last accessed 2022-26-05).

European Commission (2022b). Farm accountancy data network. URL: ec.europa.eu/info/food-farming-fisheries/farming/facts-and-figures/farms-farming-and-innovation/structures-and-economics/economics/fadn_-en#resultsandanalysis (last accessed: 2022-05-03).

European Court of Auditors (2011). Is agri-environment support well-designed and managed? Special report no. 7. Luxembourg.

European Court of Auditors (2021). Common Agricultural Policy and climate: Half of EU climate spending but farm emissions are not decreasing. Special report no. 16. Luxembourg.

European Environment Agency (2018). Environmental indicator report 2018: In support to the monitoring of the Seventh Environment Action Programme. EEA Report: No 19/2018. Denmark.

European Environment Agency (2019a). *Annual European Union Greenhouse Gas Inventory 1990–2009 and Inventory Report 2019*. Technical Report 2.

European Environment Agency (2019b). The European environment - state and outlook 2020: Knowledge for transition to a sustainable Europe. URL: https://data.europa.eu/doi/10.2800/96749 (last accessed: 2020-08-08).

European Statistical Office (2018). Agriculture - greenhouse gas emission statistics. Statistics Explained. URL: https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Archive:Agriculture_-_greenhouse_gas_emission_statistics (last accessed: 2020-08-08).

European Statistical Office (2020). Agriculture - Farm structure. URL: http://ec.europa.eu/eurostat/data/database (last accessed: 2020-05-12).

Ewald, J. A., Wheatley, C. J., Aebischer, N. J., Moreby, S. J., Duffield, S. J., Crick, H. Q. P., & Morecroft, M. B. (2015). Influences of extreme weather, climate and pesticide use on invertebrates in cereal fields over 42 years. *Global change biology*, 21(11), 3931–3950.

FAO (2017). *The future of food and agriculture – Trends and challenges*. Rome.

FAOSTAT (2022a). Annual population. URL: www.fao.org/faostat/en/#data/OA (last accessed: 2022-04-27).

FAOSTAT (2022b). Climate Change: Emissions. URL: www.fao.org/faostat/en/#data (last accessed: 2022-04-27).

FAOSTAT (2022c). Crops and livestock products. URL: www.fao.org/faostat/en/#data/QCL (last accessed: 2022-04-27).

FAOSTAT (2022d). Land, Inputs and Sustainability. URL: www.fao.org/faostat/en/#data (last accessed: 2022-04-27).

Färe, R., Grosskopf, S., & Pasurka, C. (1986). Effects on relative efficiency in electric power generation due to environmental controls. *Resources and Energy*, 8(2), 167–184.

Färe, R., Grosskopf, S., & Whittaker, G. (2013). Directional output distance functions: Endogenous directions based on exogenous normalization constraints. *Journal of Productivity Analysis*, 40(3), 267–269.

Faria, N. & Morales, M. B. (2020). Farmland management regulates ecosystem services in Mediterranean drylands: Assessing the sustainability of agri-environmental payments for bird conservation. *Journal for Nature Conservation*, 58(September), 125913.

Farr, M., Eagle, L., & Hay, R. (2018). Key determinants of pro-environmental behaviour of land managers in the agricultural sector: literature review. *Report to the National Environmental Science Program*.

Farrell, M. H., Liang, T., & Misra, S. (2021). Deep Neural Networks for Estimation and Inference. *Econometrica*, 89(1), 181–213.

Featherstone, A. M. & Goodwin, B. K. (1993). Factors Influencing a Farmer's Decision to Invest in Long-Term Conservation Improvements. *Land Economics*, 69(1), 67.

Ferraro, P. J. (2008). Asymmetric information and contract design for payments for environmental services. *Ecological Economics*, 65(4), 810–821.

Field, B. C. & Field, M. K. (2009). *Environmental Economics : An Introduction*. Boston: McGraw-Hill Irwin, 5th edition.

Filippini, M. & Greene, W. (2016). Persistent and transient productive inefficiency: a maximum simulated likelihood approach. *Journal of Productivity Analysis*, 45(2), 187–196.

Førsund, F. R. (2018). Multi-equation modelling of desirable and undesirable outputs satisfying the materials balance. *Empirical Economics*, 54(1), 67–99.

Frey, G. E., Mercer, D. E., Cubbage, F. W., & Abt, R. C. (2013). A real options model to assess the role of flexibility in forestry and agroforestry adoption and disadoption in the Lower Mississippi Alluvial Valley. *Agricultural Economics*, 44(1), 73–91.

Frick, F. & Sauer, J. (2018). Deregulation and Productivity: Empirical Evidence on Dairy Production. *American Journal of Agricultural Economics*, 100(1), 354–378.

Früh-Müller, A., Bach, M., Breuer, L., Hotes, S., Koellner, T., Krippes, C., & Wolters, V. (2019). The use of agri-environmental measures to address environmental pressures in Germany: Spatial mismatches and options for improvement. *Land Use Policy*, 84(October 2018), 347–362.

Fuentes-Montemayor, E., Goulson, D., & Park, K. J. (2011). The effectiveness of agri-environment schemes for the conservation of farmland moths: Assessing the importance of a landscape-scale management approach. *Journal of Applied Ecology*, 48(3), 532–542.

Gadanakis, Y., Bennett, R., Park, J., & Areal, F. J. (2015). Evaluating the Sustainable Intensification of arable farms. *Journal of Environmental Management*, 150, 288–298.

Galdeano-Gómez, E., Aznar-Sánchez, J. A., Pérez-Mesa, J. C., & Piedra-Muñoz, L. (2017). Exploring Synergies Among Agricultural Sustainability Dimensions: An Empirical Study on Farming System in Almería (Southeast Spain). *Ecological Economics*, 140, 99–109.

Gatto, P., Mozzato, D., & Defrancesco, E. (2019). Analysing the role of factors affecting farmers' decisions to continue with agri-environmental schemes from a temporal perspective. *Environmental Science and Policy*, 92(December 2018), 237–244.

Gerber, P., Steinfeld, H., Henderson, B., Mottet, A., Opio, C., Dijkman, J., Falcucci, A., & Tempio, G. (2013). *Tackling climate change through livestock – A global assessment of emissions and mitigation opportunities*. Rome: Food and Agriculture Organization of the United Nations (FAO).

Gillich, C., Narjes, M., Krimly, T., & Lippert, C. (2019). Combining choice modeling estimates and stochastic simulations to assess the potential of new crops—The case of lignocellulosic perennials in Southwestern Germany. *GCB Bioenergy*, 11(1), 289–303.

Girard, J., Delacote, P., & Leblois, A. (2021). Agricultural households' adaptation to weather shocks in Sub-Saharan Africa: Implications for land-use change and deforestation. *Environment and Development Economics*, 26(5-6), 538–560.

Godoy-Durán, Á., Galdeano- Gómez, E., Pérez-Mesa, J. C., & Piedra-Muñoz, L. (2017). Assessing eco-efficiency and the determinants of horticultural family-farming in southeast Spain. *Journal of Environmental Management*, 204, 594–604.

Gomes, L. C., Bianchi, F. J., Cardoso, I. M., Fernandes, R. B., Filho, E. I., & Schulte, R. P. (2020). Agroforestry systems can mitigate the impacts of climate change on coffee production: A spatially explicit assessment in Brazil. *Agriculture, Ecosystems and Environment*, 294(January), 106858.

Gómez-Calvet, R., Conesa, D., Gómez-Calvet, A. R., & Tortosa-Ausina, E. (2016). On the dynamics of eco-efficiency performance in the European Union. *Computers and Operations Research*, 66, 336–350.

Gómez-Limón, J. A., Gutiérrez-Martín, C., & Villanueva, A. J. (2019a). Optimal Design of Agri-environmental Schemes under Asymmetric Information for Improving Farmland Biodiversity. *Journal of Agricultural Economics*, 70(1), 153–177.

Gómez-Limón, J. A., Gutiérrez-Martín, C., & Villanueva, A. J. (2019b). Optimal Design of Agri-environmental Schemes under Asymmetric Information for Improving Farmland Biodiversity. *Journal of Agricultural Economics*, 70(1), 153–177.

Gonzalez-Mejia, A., Styles, D., Wilson, P., & Gibbons, J. (2018). Metrics and methods for characterizing dairy farm intensification using farm survey data. *PloS one*, 13(5), e0195286.

Gosling, E., Knoke, T., Reith, E., Reyes Cáceres, A., & Paul, C. (2021). Which Socio-economic Conditions Drive the Selection of Agroforestry at the Forest Frontier? *Environmental Management*, 67(6), 1119–1136.

Gosling, E., Reith, E., Knoke, T., & Paul, C. (2020). A goal programming approach to evaluate agroforestry systems in Eastern Panama. *Journal of Environmental Management*, 261(January), 110248.

Gossner, M. M., Lewinsohn, T. M., Kahl, T., Grassein, F., Boch, S., Prati, D., Birkhofer, K., Renner, S. C., Sikorski, J., Wubet, T., Arndt, H., Baumgartner, V., Blaser, S., Blüthgen, N., Börschig, C., Buscot, F., Diekötter, T., Jorge, L. R., Jung, K., Keyel, A. C., Klein, A.-M., Klemmer, S., Krauss, J., Lange, M., Müller, J., Overmann, J., Pašalić, E., Penone, C., Perović, D. J., Purschke, O., Schall, P., Socher, S. A., Sonnemann, I., Tschapka, M., Tscharntke, T., Türke, M., Venter, P. C., Weiner, C. N., Werner, M., Wolters, V., Wurst, S., Westphal, C., Fischer, M., Weisser, W. W., & Allan, E. (2016). Land-use intensification causes multitrophic homogenization of grassland communities. *Nature*, 540(7632), 266–269.

Granlund, K., Räike, A., Ekholm, P., Rankinen, K., & Rekolainen, S. (2005). Assessment of water protection targets for agricultural nutrient loading in Finland. *Journal of Hydrology*, 304(1-4), 251–260.

Grassauer, F., Herndl, M., Nemecek, T., Guggenberger, T., Fritz, C., Steinwidder, A., & Zollitsch, W. (2021). Eco-efficiency of farms considering multiple functions of agriculture: Concept and results from Austrian farms. *Journal of Cleaner Production*, 297, 126662.

Greene, W. (2005). Reconsidering heterogeneity in panel data estimators of the stochastic frontier model. *Journal of Econometrics*, 126(2), 269–303.

Grzelak, A., Guth, M., Matuszczak, A., Czyżewski, B., & Brelik, A. (2019). Approaching the environmental sustainable value in agriculture: How factor endowments foster the eco-efficiency. *Journal of Cleaner Production*, 241.

Guerci, M., Bava, L., Zucali, M., Sandrucci, A., Penati, C., & Tamburini, A. (2013). Effect of farming strategies on environmental impact of intensive dairy farms in Italy. *Journal of Dairy Research*, 80(3), 300–308.

Haenel, H.-D., Rösemann, C., Dämmgen, U., Döring, U., Wulf, S., Eurich-Menden, B., Freibauer, A., Döhler, H., Schreiner, C., & Osterburg, B. (2018). *Calculations of gaseous and particulate emissions from German agriculture 1990 – 2016 : Report on methods and data (RMD) Submission 2018*. Braunschweig: Johann Heinrich von Thuenen-Institut, Thuenen Rep 57.

Hansson, H., Manevska-Tasevska, G., & Asmild, M. (2020). Rationalising inefficiency in agricultural production - The case of Swedish dairy agriculture. *European Review of Agricultural Economics*, 47(1), 1–24.

Haqiqi, I., Grogan, D. S., Hertel, T. W., & Schlenker, W. (2021). Quantifying the impacts of compound extremes on agriculture. *Hydrology and Earth System Sciences*, 25(2), 551–564.

Hashida, Y. & Lewis, D. J. (2022). Estimating welfare impacts of climate change using a discrete-choice model of land management: An application to western U.S. forestry. *Resource and Energy Economics*, 68, 101295.

Hasler, B., Termansen, M., Nielsen, H. Ø., Daugbjerg, C., Wunder, S., & Latacz-Lohmann, U. (2022). European Agri-environmental Policy: Evolution, Effectiveness, and Challenges. *Review of Environmental Economics and Policy*, 16(1), 105–125.

Hauk, S., Knoke, T., & Wittkopf, S. (2014). Economic evaluation of short rotation coppice systems for energy from biomass - A review. *Renewable and Sustainable Energy Reviews*, 29, 435–448.

He, X. & Chen, Z. (2022). Weather, cropland expansion, and deforestation in Ethiopia. *Journal of Environmental Economics and Management*, 111, 102586.

Heiler, P. & Knaus, M. C. (2021). Effect or Treatment Heterogeneity? Policy Evaluation with Aggregated and Disaggregated Treatments. URL: http://arxiv.org/abs/2110.01427 (last accessed: 2021-12-1).

Heinz, S., Mayer, F., & Kuhn, G. (2015). *Grünlandmonitoring Bayern - Evaluierung von Agrarumweltmaßnahmen*. Freising-Weihenstephan.

Henningsen, A. (2019). Introduction to Econometric Production Analysis with R. Collection of Lecture Notes. 3rd Draft Version. Department of Food and Resource Economics, University of Copenhagen. URL: http://leanpub.com/ProdEconR/ (last accessed: 2022-6-3).

Henningsen, A. & Henning, C. H. (2009). Imposing regional monotonicity on translog stochastic production frontiers with a simple three-step procedure. *Journal of Productivity Analysis*, 32(3), 217–229.

Hensher, D. A., Rose, J. M., & Greene, W. H. (2015). *Applied Choice Analysis*. Cambridge: Cambridge University Press.

Hernández-Morcillo, M., Burgess, P., Mirck, J., Pantera, A., & Plieninger, T. (2018). Scanning agroforestry-based solutions for climate change mitigation and adaptation in Europe. *Environmental Science and Policy*, 80(November 2017), 44–52.

Herrero, M., Henderson, B., Havlík, P., Thornton, P. K., Conant, R. T., Smith, P., Wirsenius, S., Hristov, A. N., Gerber, P., Gill, M., Butterbach-Bahl, K., Valin, H., Garnett, T., & Stehfest, E. (2016). Greenhouse gas mitigation potentials in the livestock sector. *Nature Climate Change*, 6(5), 452–461.

Hess, S. & Rose, J. M. (2012). Can scale and coefficient heterogeneity be separated in random coefficients models? *Transportation*, 39(6), 1225–1239.

Hess, S. & Train, K. (2017). Correlation and scale in mixed logit models. *Journal of Choice Modelling*, 23, 1–8.

Holland, P. W. (1986). Statistics and Causal Inference. *Journal of the American Statistical Association*, 81(396), 945–960.

Hotz, V. J., Imbens, G. W., & Mortimer, J. H. (2005). Predicting the efficacy of future training programs using past experiences at other locations. *Journal of Econometrics*, 125(1-2 SPEC. ISS.), 241–270.

Huang, C. J., Huang, T.-H., & Liu, N.-H. (2014). A new approach to estimating the metafrontier production function based on a stochastic frontier framework. *Journal of Productivity Analysis*, 42(3), 241–254.

Huber, R., Snell, R., Monin, F., Sibyl, H. B., Schmatz, D., & Finger, R. (2017). Interaction effects of targeted agri-environmental payments on non-marketed goods and services under climate change in a mountain region. *Land Use Policy*, 66(April), 49–60.

Huber, W. (2020). Generation of a random variable with fixed covariance structure. URL: stats.stackexchange.com/questions/444039/whuber-s-generation-of-a-random-variable-with-fixed-covariance-structure (last accessed: 28-10-2020).

Hünermund, P., Louw, B., & Caspi, I. (2021). Double Machine Learning and Automated Confounder Selection – A Cautionary Tale. URL: http://arxiv.org/abs/2108.11294 (last accessed: 2021-1-22).

Huppes, G. & Ishikawa, M. (2005). A Framework for Quantified Eco-efficiency Analysis. *Journal of Industrial Ecology*, 9(4), 25–41.

Iacobucci, D., Schneider, M. J., Popovich, D. L., & Bakamitsos, G. A. (2016). Mean centering helps alleviate "micro" but not "macro" multicollinearity. *Behavior Research Methods*, 48(4), 1308–1317.

Iglesias, A., Quiroga, S., Moneo, M., & Garrote, L. (2012). From climate change impacts to the development of adaptation strategies: Challenges for agriculture in Europe. *Climatic Change*, 112(1), 143–168.

International Fertilizer Industry Association (IFA) (1992). *World Fertilizer Use Manual*. Paris: IFA, International Fertilizer Industry Association.

IPCC (2006). *2006 IPCC Guidelines for National Greenhouse Gas Inventories*. Japan: National Greenhouse Gas Inventories Programme.

IPCC (2007). *Climate Change 2007: Impacts, Adaptation and Vulnerability - Contribution of Working Group II to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge: Cambridge University Press.

IPCC (2013). *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge and New York: Cambridge University Press.

IPCC (2014). Climate Change 2014, Mitigation of Climate Change. In T. Z. Edenhofer, O., R. Pichs-Madruga, Y. Sokona, E. Farahani, S. Kadner, K. Seyboth, A. Adler, I. Baum, S. Brunner, P. Eickemeier, B. Kriemann, J. Savolainen, S. Schlömer, C. von Stechow & J. Minx (Eds.), *Working Group III Contribution to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, Summary for Policymakers (pp.$\tilde{3}$5). Cambridge and New York: Cambridge University Press.

IPCC (2021). *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge: Cambridge University Press.

Iribarren, D., Hospido, A., Moreira, M. T., & Feijoo, G. (2011). Benchmarking environmental and operational parameters through eco-efficiency criteria for dairy farms. *Science of the Total Environment*, 409(10), 1786–1798.

ISO 14044:2006 (2006). Environmental Management – Life Cycle Assessment – Requirements and Guidelines. Geneva: International Organization for Standardization.

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2021). *An Introduction to Statistical Learning*, volume 103 of *Springer Texts in Statistics*. New York: Springer, second edition.

Kaligarič, M., Čuš, J., Škornik, S., & Ivajnšič, D. (2019). The failure of agri-environment measures to promote and conserve grassland biodiversity in Slovenia. *Land Use Policy*, 80, 127–134.

Kallus, N., Mao, X., & Udell, M. (2018). Causal Inference with Noisy and Missing Covariates via Matrix Factorization. URL: http://arxiv.org/abs/1806.00811 last accessed: 2021-12-1).

Kamilaris, A., Kartakoullis, A., & Prenafeta-Boldú, F. X. (2017). A review on the practice of big data analysis in agriculture. *Computers and Electronics in Agriculture*, 143(January), 23–37.

Katsanevakis, S., Bogucarskis, K., Gatto, F., Vandekerkhove, J., Deriu, I., & Cardoso, A. C. (2012). Building the European Alien Species Information Network (EASIN): a novel approach for the exploration of distributed alien species data. *BioInvasions Records*, 1(4), 235–245.

Kawasaki, K. & Uchida, S. (2016). Quality Matters More Than Quantity: Asymmetric Temperature Effects on Crop Yield and Quality Grade. *American Journal of Agricultural Economics*, 98(4), 1195–1209.

Keane, M. & Neal, T. (2020). Climate change and U.S. agriculture: Accounting for multidimensional slope heterogeneity in panel data. *Quantitative Economics*, 11(4), 1391–1429.

Kellermann, M. A. (2015). Total Factor Productivity Decomposition and Unobserved Heterogeneity in Stochastic Frontier Models. *Agricultural and Resource Economics Review*, 44(1), 124–148.

King, G. & Zeng, L. (2006). The dangers of extreme counterfactuals. *Political Analysis*, 14(2), 131–159.

Kleijn, D., Berendse, F., Smit, R., Gilissen, N., Smit, J., Brak, B., & Groeneveld, R. (2004). Ecological Effectiveness of Agri-Environment Schemes in Different Agricultural Landscapes in The Netherlands. *Conservation Biology*, 18(3), 775–786.

Klein, T., Holzkämper, A., Calanca, P., & Fuhrer, J. (2014). Adaptation options under climate change for multifunctional agriculture: A simulation study for western Switzerland. *Regional Environmental Change*, 14(1), 167–184.

Knoke, T., Paul, C., Härtl, F., Castro, L. M., Calvas, B., & Hildebrandt, P. (2015). Optimizing agricultural land-use portfolios with scarce data-A non-stochastic model. *Ecological Economics*, 120, 250–259.

Knudson, W. A. (2009). The Environment, Energy, and the Tinbergen Rule. *Bulletin of Science, Technology & Society*, 29(4), 308–312.

Kopittke, P. M., Menzies, N. W., Wang, P., McKenna, B. A., & Lombi, E. (2019). Soil and the intensification of agriculture for global food security. *Environment International*, 132(August), 105078.

Kortelainen, M. (2008). Dynamic environmental performance analysis: A Malmquist index approach. *Ecological Economics*, 64(4), 701–715.

Kragt, M. E. & Robertson, M. J. (2014). Quantifying ecosystem services trade-offs from agricultural practices. *Ecological Economics*, 102, 147–157.

Kuersten, E. & Burschel, P. (1993). CO2-mitigation by agroforestry. *Water, Air, & Soil Pollution*, 70(1-4), 533–544.

Kuhfuss, L., Préget, R., Thoyer, S., & Hanley, N. (2016). Nudging farmers to enrol land into agri-environmental schemes: The role of a collective bonus. *European Review of Agricultural Economics*, 43(4), 609–636.

Kuhfuss, L. & Subervie, J. (2018). Do European Agri-environment Measures Help Reduce Herbicide Use? Evidence From Viticulture in France. *Ecological Economics*, 149(October 2017), 202–211.

Kumbhakar, S. C. & Bokusheva, R. (2009). Modelling farm production decisions under an expenditure constraint. *European Review of Agricultural Economics*, 36(3), 343–367.

Kumbhakar, S. C., Lien, G., & Hardaker, J. B. (2014). Technical efficiency in competing panel data models: A study of Norwegian grain farming. *Journal of Productivity Analysis*, 41(2), 321–337.

Kumbhakar, S. C. & Lovell, C. A. K. (2000). *Stochastic Frontier Analysis*. Cambridge, UK: Cambridge University Press.

Kumbhakar, S. C. & Tsionas, E. G. (2016). The good, the bad and the technology: Endogeneity in environmental production models. *Journal of Econometrics*, 190(2), 315–327.

Kumbhakar, S. C., Wang, H.-J., & Horncastle, A. P. (2015). *A practitioner's guide to stochastic frontier analysis using Stata*. New York: Cambridge University Press.

Künzel, S. R., Sekhon, J. S., Bickel, P. J., & Yu, B. (2019). Metalearners for estimating heterogeneous treatment effects using machine learning. *Proceedings of the National Academy of Sciences of the United States of America*, 116(10), 4156–4165.

Kuosmanen, T. & Kortelainen, M. (2005). Measuring Eco-efficiency of Production with Data. *Journal of Industrial Ecology*, 9(4), 59–72.

Lai, H.-p. & Kumbhakar, S. C. (2021). A panel frontier system model with good and bad outputs and endogenous treatment decision. *Economics Letters*, 198, 109649.

Lamkowsky, M., Oenema, O., Meuwissen, M. P. M., & Ang, F. (2021). Closing productivity gaps among Dutch dairy farms can boost profit and reduce nitrogen pollution. *Environmental Research Letters*, 16(12), 124003.

Lancaster, K. J. (1966). A New Approach to Consumer Theory. *Journal of Political Economy*, 74(2), 132–157.

Landini, F., Arrighetti, A., & Bartoloni, E. (2020). The sources of heterogeneity in firm performance: Lessons from Italy. *Cambridge Journal of Economics*, 44(3), 527–558.

Landry, C. E. & List, J. A. (2007). Using ex ante approaches to obtain credible signals for value in contingent markets: Evidence from the field. *American Journal of Agricultural Economics*, 89(2), 420–429.

Landwirtschaftskammer Niedersachsen (2022). Die neue GAP ab 2023 - eine ökonomische Optimierung der Anträge wird wichtiger! URL: www.lwk-niedersachsen.de/lwk/news/38437 (last accessed 2022-05-06).

Langenberg, J., Rauert, M., & Theuvsen, L. (2018). Einstellungen landwirtschaftlicher Betriebsleiter und weiterer in Deutschland: Eine empirische Analyse anhand von Stakeholder zu silvoarablen Alley-Cropping-Agroforstsystemen Experteninterviews. *Berichte über Landwirtschaft*, 96(2), 1–35.

Langenberg, J. & Theuvsen, L. (2018). Agroforstwirtschaft in Deutschland : Alley-Cropping-Systeme aus ökonomischer Perspektive. 70(4).

Langpap, C., Hascic, I., & Wu, J. (2008). Protecting watershed ecosystems through targeted local land use policies. *American Journal of Agricultural Economics*, 90(3), 684–700.

Lasch, P., Kollas, C., Rock, J., & Suckow, F. (2010). Potentials and impacts of short-rotation coppice plantation with aspen in Eastern Germany under conditions of climate change. *Regional Environmental Change*, 10(2), 83–94.

Lasco, R. D., Delfino, R. J. P., Catacutan, D. C., Simelton, E. S., & Wilson, D. M. (2014). Climate risk adaptation by smallholder farmers: The roles of trees and agroforestry. *Current Opinion in Environmental Sustainability*, 6(1), 83–88.

Lasco, R. D., Espaldon, M. L. O., & Habito, C. M. D. (2016). Smallholder farmers' perceptions of climate change and the roles of trees and agroforestry in climate risk adaptation: evidence from Bohol, Philippines. *Agroforestry Systems*, 90(3), 521–540.

Latacz-Lohmann, U. & Breustedt, G. (2019). Using choice experiments to improve the design of agri-environmental schemes. *European Review of Agricultural Economics*, 46(3), 495–528.

Latacz-Lohmann, U., Termansen, M., & Nguyen, C. (2022). The New Eco-Schemes: Navigating a Narrow Fairway. *EuroChoices*, 0(0), 1–7.

Latacz-Lohmann, U. & Van der Hamsvoort, C. (1997). Auctioning conservation contracts: A theoretical analysis and an application. *American Journal of Agricultural Economics*, 79(2), 407–418.

Lauwers, L. (2009). Justifying the incorporation of the materials balance principle into frontier-based eco-efficiency models. *Ecological Economics*, 68(6), 1605–1614.

Layton, D. F. & Siikamäki, J. (2009). Payments for ecosystem services programs: Predicting landowner enrollment and opportunity cost using a beta-binomial model. *Environmental and Resource Economics*, 44(3), 415–439.

Le, S., Jeffrey, S., & An, H. (2020). Greenhouse Gas Emissions and Technical Efficiency in Alberta Dairy Production: What Are the Trade-Offs? *Journal of Agricultural and Applied Economics*, 52(2), 177–193.

Leakey, R. R. (2017). Definition of Agroforestry Revisited. In *Multifunctional Agriculture*, number January 1996 (pp. 5–6). Elsevier.

Lechner, M. (2019). Modified Causal Forests for Estimating Heterogeneous Causal Effects. URL: http://arxiv.org/abs/1812.09487 (last accessed: 2021-12-04).

Leonhardt, H., Braito, M., & Uehleke, R. (2022). Combining the best of two methodological worlds? Integrating Q methodology-based farmer archetypes in a quantitative model of agri-environmental scheme uptake. *Agriculture and Human Values*, 39(1), 217–232.

Lesk, C., Rowhani, P., & Ramankutty, N. (2016). Influence of extreme weather disasters on global crop production. *Nature*, 529(7584), 84–87.

LfL (2015). *Agrarstrukturentwicklung in Bayern - IBA -Agrarstrukturbericht 2014*. Freising-Weihenstephan: Bavarian State Research Center for Agriculture.

LfL (2018). LfL Deckungsbeiträge und Kalkulationsdaten. Bavarian State Research Centre for Agriculture. URL: https://www.stmelf.bayern.de/idb/default.html (last accessed: 2022-01-03).

Lichtenberg, E. (2002). Chapter 23 Agriculture and the environment. In *Handbook of Agricultural Economics*, volume 2 (pp. 1249–1313).

Lipton, Z. C. (2018). The Mythos of Model Interpretability: In machine learning, the concept of interpretability is both important and slippery. *Queue*, 16(3), 31–57.

Liu, P. & Tian, X. (2021). Downward Hypothetical Bias in the Willingness to Accept Measure for Private Goods: Evidence from a Field Experiment. *American Journal of Agricultural Economics*, 103(5), 1679–1699.

Loiseau, E., Saikku, L., Antikainen, R., Droste, N., Hansjürgens, B., Pitkänen, K., Leskinen, P., Kuikman, P., & Thomsen, M. (2016). Green economy and related concepts: An overview. *Journal of Cleaner Production*, 139, 361–371.

Londo, M., Vleeshouwers, L., Dekker, J., & de Graaf, H. (2001). Energy farming in Dutch desiccation abatement areas: yields and benefits compared to grass cultivation. *Biomass and Bioenergy*, 20(5), 337–350.

Long, T. B., Blok, V., & Coninx, I. (2016). Barriers to the adoption and diffusion of technological innovations for climate-smart agriculture in Europe: Evidence from the Netherlands, France, Switzerland and Italy. *Journal of Cleaner Production*, 112, 9–21.

Longo, M., Dal Ferro, N., Lazzaro, B., & Morari, F. (2021). Trade-offs among ecosystem services advance the case for improved spatial targeting of agri-environmental measures. *Journal of Environmental Management*, 285(February), 112131.

Lötjönen, S., Temmes, E., & Ollikainen, M. (2020). Dairy Farm Management when Nutrient Runoff and Climate Emissions Count. *American Journal of Agricultural Economics*, 102(3), 960–981.

Louizos, C., Shalit, U., Mooij, J., Sontag, D., Zemel, R., & Welling, M. (2017). Causal effect inference with deep latent-variable models. *Proceedings of the 31st International Conference on Neural Information Processing Systems*, (pp. 6446–6456).

Louviere, J. J., Hensher, D. A., Swait, J. D., & Adamowicz, W. (2000). *Stated Choice Methods*. Cambridge, UK: Cambridge University Press.

Lovén, I. & Nordin, M. (2020). Are agri-environmental schemes boosting farm survival? AgriFood Economics Centre Working Paper 2020:2. URL: https://agrifood.se/Files/AgriFood_WP20202.pdf (last accessed: 2022-6-7).

Lundberg, S. M. & Lee, S.-I. (2017). A unified approach to interpreting model predictions. In I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, & R. Garnett (Eds.), *Advances in Neural Information Processing Systems*, volume 30 (pp. 4765–4774).

Lüttger, A. B. & Feike, T. (2018). Development of heat and drought related extreme weather events and their effect on winter wheat yields in Germany. *Theoretical and Applied Climatology*, 132(1-2), 15–29.

Lynch, J., Cain, M., Frame, D., & Pierrehumbert, R. (2021). Agriculture's Contribution to Climate Change and Role in Mitigation Is Distinct From Predominantly Fossil CO2-Emitting Sectors. *Frontiers in Sustainable Food Systems*, 4(February), 1–9.

MacDonald, M. A., Cobbold, G., Mathews, F., Denny, M. J. H., Walker, L. K., Grice, P. V., & Anderson, G. Q. A. (2012). Effects of agri-environment management for cirl buntings on other biodiversity. *Biodiversity and Conservation*, 21(6), 1477–1492.

Mack, G., Ritzel, C., & Jan, P. (2020). Determinants for the Implementation of Action-, Result- and Multi-Actor-Oriented Agri-Environment Schemes in Switzerland. *Ecological Economics*, 176(March), 106715.

Mahmood, I., Imadi, S. R., Shazadi, K., Gul, A., & Hakeem, K. R. (2016). Effects of Pesticides on Environment. In *Plant, Soil and Microbes* (pp. 253–269). Cham: Springer International Publishing.

Mahmoud, N. & Hutchings, N. J. (2020). The advantages of using field- and farm-scale data to target agri-environmental measures—an example of afforestation. *Environmental Science and Policy*, 114(March), 14–21.

Malek, Ž., Douw, B., Van Vliet, J., Van Der Zanden, E. H., & Verburg, P. H. (2019). Local land-use decision-making in a global context. *Environmental Research Letters*, 14(8), 083006.

Malikov, E., Bokusheva, R., & Kumbhakar, S. C. (2018). A hedonic-output-index-based approach to modeling polluting technologies. *Empirical Economics*, 54(1), 287–308.

Managi, S. (2006). Are there increasing returns to pollution abatement? Empirical analytics of the Environmental Kuznets Curve in pesticides. *Ecological Economics*, 58(3), 617–636.

Managi, S., Opaluch, J. J., Jin, D., & Grigalunas, T. A. (2005). Environmental Regulations and Technological Change in the Offshore Oil and Gas Industry. *Land Economics*, 81(2), 303–319.

Mangham, L. J., Hanson, K., & McPake, B. (2009). How to do (or not to do) ... Designing a discrete choice experiment for application in a low-income country. *Health Policy and Planning*, 24(2), 151–158.

Mann, M. E., Rahmstorf, S., Kornhuber, K., Steinman, B. A., Miller, S. K., Petri, S., & Coumou, D. (2018). Projected changes in persistent extreme summer weather events : The role of quasi-resonant amplification. *Science advances*, 4(October), 1–9.

Mariel, P., Hoyos, D., Meyerhoff, J., Czajkowski, M., Dekker, T., Glenk, K., Jacobsen, J. B., Liebe, U., Olsen, S. B., Sagebiel, J., & Thiene, M. (2021). *Environmental Valuation with Discrete Choice Experiments*. Springer Briefs in Economics. Cham: Springer International Publishing.

Mariel, P. & Meyerhoff, J. (2018). A More Flexible Model or Simply More Effort? On the Use of Correlated Random Parameters in Applied Choice Studies. *Ecological Economics*, 154(July), 419–429.

Martinez Cillero, M., Thorne, F., Wallace, M., & Breen, J. (2019). Technology heterogeneity and policy change in farm-level efficiency analysis: an application to the Irish beef sector. *European Review of Agricultural Economics*, 46(2), 193–214.

Martinez Cillero, M., Thorne, F., Wallace, M., Breen, J., & Hennessy, T. (2018). The Effects of Direct Payments on Technical Efficiency of Irish Beef Farms: A Stochastic Frontier Analysis. *Journal of Agricultural Economics*, 69(3), 669–687.

Martinsson, E. & Hansson, H. (2021). Adjusting eco-efficiency to greenhouse gas emissions targets at farm level – The case of Swedish dairy farms. *Journal of Environmental Management*, 287(November 2020), 112313.

Mattia, C. M., Lovell, S. T., & Davis, A. (2018). Identifying barriers and motivators for adoption of multifunctional perennial cropping systems by landowners in the Upper Sangamon River Watershed, Illinois. *Agroforestry Systems*, 92(5), 1155–1169.

Matzdorf, B., Kaiser, T., & Rohner, M. S. (2008). Developing biodiversity indicator to design efficient agri-environmental schemes for extensively used grassland. *Ecological Indicators*, 8(3), 256–269.

McFadden, D. (1973). Conditional logit analysis of qualitative choice behavior.

McGinty, M. M., Swisher, M. E., & Alavalapati, J. (2008). Agroforestry adoption and maintenance: Self-efficacy, attitudes and socio-economic factors. *Agroforestry Systems*, 73(2), 99–108.

McGurk, E., Hynes, S., & Thorne, F. (2020). Participation in agri-environmental schemes: A contingent valuation study of farmers in Ireland. *Journal of Environmental Management*, 262(February 2019), 110243.

Meeusen, W. & van den Broeck, J. (1977). Technical efficiency and dimension of the firm: Some results on the use of frontier production functions. *Empirical Economics*, 2(2), 109–122.

Meijer, S. S., Catacutan, D., Sileshi, G. W., & Nieuwenhuis, M. (2015). Tree planting by smallholder farmers in Malawi: Using the theory of planned behaviour to examine the relationship between attitudes and behaviour. *Journal of Environmental Psychology*, 43, 1–12.

Menapace, L., Colson, G., & Raffaelli, R. (2013). Risk aversion, subjective beliefs, and farmer risk management strategies. *American Journal of Agricultural Economics*, 95(2), 384–389.

Mendelsohn, R., Nordhaus, W. D., & Shaw, D. (1994). American Economic Association The Impact of Global Warming on Agriculture : A Ricardian Analysis. *The American Economic Review*, 84(4), 753–771.

Mennig, P. & Sauer, J. (2020). The impact of agri-environment schemes on farm productivity: a DID-matching approach. *European Review of Agricultural Economics*, 47(3), 1045–1093.

Mercer, E. & Snook, A. (2005). Analyzing Ex-Ante Agroforestry Adoption Decisions with Attribute-Based Choice Experiments. In *Valuing Agroforestry Systems* (pp. 237–256). Dordrecht: Kluwer Academic Publishers.

Meuwissen, M. P., Feindt, P. H., Spiegel, A., Termeer, C. J., Mathijs, E., de Mey, Y., Finger, R., Balmann, A., Wauters, E., Urquhart, J., Vigani, M., Zawalińska, K., Herrera, H., Nicholas-Davies, P., Hansson, H., Paas, W., Slijper, T., Coopmans, I., Vroege, W., Ciechomska, A., Accatino, F., Kopainsky, B., Poortvliet, P. M., Candel, J. J., Maye, D., Severini, S., Senni, S., Soriano, B., Lagerkvist, C. J., Peneva, M., Gavrilescu, C., & Reidsma, P. (2019). A framework to assess the resilience of farming systems. *Agricultural Systems*, 176(January), 102656.

Mfitumukiza, D., Barasa, B., & Ingrid, A. (2017). Determinants of agroforestry adoption as an adaptation means to drought among smallholder farmers in Nakasongola District, Central Uganda. *African Journal of Agricultural Research*, 12(23), 2024–2035.

Miller, S. (2020). Causal forest estimation of heterogeneous and time-varying environmental policy effects. *Journal of Environmental Economics and Management*, 103, 102337.

Minviel, J. J. & Sipiläinen, T. (2018). Dynamic stochastic analysis of the farm subsidy-efficiency link: evidence from France. *Journal of Productivity Analysis*, 50(1-2), 41–54.

Möhring, N., Gaba, S., & Finger, R. (2019). Quantity based indicators fail to identify extreme pesticide risks. *Science of The Total Environment*, 646, 503–523.

Molnar, C. (2018). iml: An R package for Interpretable Machine Learning. *Journal of Open Source Software*, 3(26), 786.

Molnar, C. (2019). Interpretable Machine Learning. A Guide for Making Black Box Models Explainable. URL: https://christophm.github.io/interpretable-ml-book/ (last accessed: 2020-11-11).

Montgomery, J. M., Nyhan, B., & Torres, M. (2018). How conditioning on post-treatment variables can ruin your experiment and what to do about it. *American Journal of Political Science*, 62(3), 760–775.

Morton, L. W., Hobbs, J., Arbuckle, J. G., & Loy, A. (2015). Upper Midwest Climate Variations: Farmer Responses to Excess Water Risks. *Journal of Environmental Quality*, 44(3), 810–822.

Motoaki, Y. & Daziano, R. A. (2015). A hybrid-choice latent-class model for the analysis of the effects of weather on cycling demand. *Transportation Research Part A: Policy and Practice*, 75(2015), 217–230.

Mu, W., Kanellopoulos, A., van Middelaar, C. E., Stilmant, D., & Bloemhof, J. M. (2018). Assessing the impact of uncertainty on benchmarking the eco-efficiency of dairy farming using fuzzy data envelopment analysis. *Journal of Cleaner Production*, 189, 709–717.

Muffler, K. & Ulber, R. (2008). Use of renewable raw materials in the chemical industry - Beyond sugar and starch. *Chemical Engineering and Technology*, 31(5), 638–646.

Mullally, C. & Chakravarty, S. (2018). Are matching funds for smallholder irrigation money well spent? *Food Policy*, 76(July 2017), 70–80.

Murphy, J. J., Stevens, T. H., & Weatherhead, D. (2005). Is cheap talk effective at eliminating hypothetical bias in a provision point mechanism? *Environmental and Resource Economics*, 30(3), 327–343.

Murty, S., Robert Russell, R., & Levkoff, S. B. (2012). On modeling pollution-generating technologies. *Journal of Environmental Economics and Management*, 64(1), 117–135.

Musshoff, O. (2012). Growing short rotation coppice on agricultural land in Germany: A Real Options Approach. *Biomass and Bioenergy*, 41, 73–85.

Nair, P. K. R. (1985). Classification of agroforestry systems. *Agroforestry Systems*, 3(2), 97–128.

Nerlove, M. (1958). Adaptive Expectations and Cobweb Phenomena. *The Quarterly Journal of Economics*, 72(2), 227.

Neyman, J. (1923). Sur les Applications de la Théorie des Probabilités aux Experiences Agricoles: Essai des Principes. *Roczniki Nauk Rolniczych*, 10, 1–51.

Njuki, E., Bravo-Ureta, B. E., & Cabrera, V. E. (2020). Climatic effects and total factor productivity: Econometric evidence for Wisconsin dairy farms. *European Review of Agricultural Economics*, 47(3), 1276–1301.

Njuki, E., Bravo-Ureta, B. E., & Mukherjee, D. (2016). The good and the bad: Environmental efficiency in northeastern U.S. dairy farming. *Agricultural and Resource Economics Review*, 45(1), 22–43.

Njuki, E., Bravo-Ureta, B. E., & O'Donnell, C. J. (2018). A new look at the decomposition of agricultural productivity growth incorporating weather effects. *PLoS ONE*, 13(2), 1–21.

Njuki, E., Bravo-Ureta, B. E., & O'Donnell, C. J. (2019). Decomposing agricultural productivity growth using a random-parameters stochastic production frontier. *Empirical Economics*, 57(3), 839–860.

Norris, K. (2008). Agriculture and biodiversity conservation: opportunity knocks. *Conservation Letters*, 1(1), 2–11.

O'Brien, D., Shalloo, L., Patton, J., Buckley, F., Grainger, C., & Wallace, M. (2012). A life cycle assessment of seasonal grass-based and confinement dairy farms. *Agricultural Systems*, 107, 33–46.

O'Donnell, C. J. (2012). An aggregate quantity framework for measuring and decomposing productivity change. *Journal of Productivity Analysis*, 38(3), 255–272.

O'Donnell, C. J. (2016). Using information about technologies, markets and firm behaviour to decompose a proper productivity index. *Journal of Econometrics*, 190(2), 328–340.

O'Donnell, C. J., Rao, D. S., & Battese, G. E. (2008). Metafrontier frameworks for the study of firm-level efficiencies and technology ratios. *Empirical Economics*, 34(2), 231–255.

OECD (2019). *Innovation, Productivity and Sustainability in Food and Agriculture*. Paris: OECD Publishing.

OECD (2020). *Strengthening Agricultural Resilience in the Face of Multiple Risks*. Paris: OECD Publishing.

OECD (2022). Agri-Environmental other indicators: Farm Birds Index. URL: stats.oecd.org/Index.aspx?QueryId=77269 (last accessed: 2022-04-27).

Oelbermann, M., Paul Voroney, R., & Gordon, A. M. (2004). Carbon sequestration in tropical and temperate agroforestry systems: A review with examples from Costa Rica and southern Canada. *Agriculture, Ecosystems and Environment*, 104(3), 359–377.

Ogunbode, C. A., Demski, C., Capstick, S. B., & Sposato, R. G. (2019). Attribution matters: Revisiting the link between extreme weather experience and climate change mitigation responses. *Global Environmental Change*, 54, 31–39.

Olesen, J. E., Schelde, K., Weiske, A., Weisbjerg, M. R., Asman, W. A., & Djurhuus, J. (2006). Modelling greenhouse gas emissions from European conventional and organic dairy farms. *Agriculture, Ecosystems and Environment*, 112(2-3), 207–220.

Olesen, J. E., Trnka, M., Kersebaum, K. C., Skjelvåg, A. O., Seguin, B., Peltonen-Sainio, P., Rossi, F., Kozyra, J., & Micale, F. (2011). Impacts and adaptation of European crop production systems to climate change. *European Journal of Agronomy*, 34(2), 96–112.

Oprescu, M., Syrgkanis, V., & Wu, Z. S. (2018). Orthogonal Random Forest for Causal Inference. URL: http://arxiv.org/abs/1806.03467 (last accessed 2021-10-10).

Orea, L. (2002). Parametric decomposition of a generalized Malmquist productivity index. *Journal of Productivity Analysis*, 18(1), 5–22.

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

Orea, L., Perez, J. A., & Roibas, D. (2015). Land Use Policy Evaluating the double effect of land fragmentation on technology choice and dairy farm productivity : A latent class model approach. *Land Use Policy*, 45, 189–198.

Orea, L. & Wall, A. (2017). A Parametric Approach to Estimating Eco-Efficiency. *Journal of Agricultural Economics*, 68(3), 901–907.

Ortiz-Bobea, A., Ault, T. R., Carrillo, C. M., Chambers, R. G., & Lobell, D. B. (2021). Anthropogenic climate change has slowed global agricultural productivity growth. *Nature Climate Change*, 11(4), 306–312.

Ortiz-Bobea, A., Knippenberg, E., & Chambers, R. G. (2018). Growing climatic sensitivity of U.S. agriculture linked to technological change and regional specialization. *Science Advances*, 4(12), 1–9.

Oviedo, J. L., Campos, P., & Caparrós, A. (2021). Contingent valuation of landowner demand for forest amenities: application in Andalusia, Spain. *European Review of Agricultural Economics*.

Pachali, M. J., Kurz, P., & Otter, T. (2020). How to generalize from a hierarchical model? *Quantitative Marketing and Economics*, 18(4), 343–380.

Páez, A. (2019). The Pragmatic Turn in Explainable Artificial Intelligence (XAI). *Minds and Machines*, 29(3), 441–459.

Panagos, P., Borrelli, P., Poesen, J., Ballabio, C., Lugato, E., Meusburger, K., Montanarella, L., & Alewell, C. (2015). The new assessment of soil loss by water erosion in Europe. *Environmental Science & Policy*, 54, 438–447.

Panagos, P., Van Liedekerke, M., Jones, A., & Montanarella, L. (2012). European Soil Data Centre: Response to European policy support and public data requirements. *Land Use Policy*, 29(2), 329–338.

Pantera, Mosquera-Losada, M. R., Herzog, F., & den Herder, M. (2021). Agroforestry and the environment. *Agroforestry Systems*, 95(5), 767–774.

Pattanayak, S. K., Mercer, D. E., Sills, E., & Yang, J. C. (2003). Taking stock of agroforestry adoption studies. *Agroforestry Systems*, 57(3), 173–186.

Paul, C., Weber, M., & Knoke, T. (2017). Agroforestry versus farm mosaic systems - Comparing land-use efficiency, economic returns and risks under climate change effects. *Science of The Total Environment*, 587-588, 22–35.

Paustian, K., Babcock, B., Hatfield, J., Kling, C., Lal, R., McCarl, B., McLaughlin, S., Mosier, A., Post, W., Rice, C., Robertson, G., Rosenberg, N., Rosenzweig, C., Schlesinger, W., & Zilberman, D. (2004). *Climate Change and Greenhouse Gas Mitigation: Challenges and Opportunities for Agriculture*, volume 141 of *Task Force Report No. 141.* Ames: Council for Agricultural Science and Technology.

Pearl, J. (2018). Theoretical Impediments to Machine Learning With Seven Sparks from the Causal Revolution. In *Proceedings of the Eleventh ACM International Conference on Web Search and Data Mining* (pp. 1–9). New York, NY, USA: ACM.

Pe'er, G., Bonn, A., Bruelheide, H., Dieker, P., Eisenhauer, N., Feindt, P. H., Hagedorn, G., Hansjürgens, B., Herzon, I., Lomba, Â., Marquard, E., Moreira, F., Nitsch, H., Oppermann, R., Perino, A., Röder, N., Schleyer, C., Schindler, S., Wolf, C., Zinngrebe, Y., & Lakner, S. (2020). Action needed for the EU Common Agricultural Policy to address sustainability challenges. *People and Nature*, 2(2), 305–316.

Pelosi, C., Goulard, M., & Balent, G. (2010). The spatial scale mismatch between ecological processes and agricultural management: Do difficulties come from underlying theoretical frameworks? *Agriculture, Ecosystems and Environment*, 139(4), 455–462.

Penn, J. & Hu, W. (2019). Cheap talk efficacy under potential and actual Hypothetical Bias: A meta-analysis. *Journal of Environmental Economics and Management*, 96, 22–35.

Penn, J. M. & Hu, W. (2021). The Extent of Hypothetical Bias in Willingness to Accept. *American Journal of Agricultural Economics*, 103(1), 126–141.

Pérez-Domínguez, I., del Prado, A., Mittenzwei, K., Hristov, J., Frank, S., Tabeau, A., Witzke, P., Havlik, P., van Meijl, H., Lynch, J., Stehfest, E., Pardo, G., Barreiro-Hurle, J., Koopman, J. F., & Sanz-Sánchez, M. J. (2021). Short- and long-term warming effects of methane may affect the cost-effectiveness of mitigation policies and benefits of low-meat diets. *Nature Food*, 2(12), 970–980.

Pérez Urdiales, M., Lansink, A. O., & Wall, A. (2016). Eco-efficiency Among Dairy Farmers: The Importance of Socio-economic Characteristics and Farmer Attitudes. *Environmental and Resource Economics*, 64(4), 559–574.

Perkins, A. J., Maggs, H. E., Watson, A., & Wilson, J. D. (2011). Adaptive management and targeting of agri-environment schemes does benefit biodiversity: A case study of the corn bunting Emberiza calandra. *Journal of Applied Ecology*, 48(3), 514–522.

Petersen, S., Blanchard, M., Chadwick, D., Del Prado, A., Edouard, N., Mosquera, J., & Sommer, S. (2013). Manure management for greenhouse gas mitigation. *Animal*, 7(s2), 266–282.

Picazo-Tadeo, A. J., Beltrán-Esteve, M., & Gómez-Limón, J. A. (2012). Assessing eco-efficiency with directional distance functions. *European Journal of Operational Research*, 220(3), 798–809.

Picazo-Tadeo, A. J., Castillo-Giménez, J., & Beltrán-Esteve, M. (2014). An intertemporal approach to measuring environmental performance with directional distance functions: Greenhouse gas emissions in the European Union. *Ecological Economics*, 100, 173–182.

Picazo-Tadeo, A. J., Gómez-Limón, J. A., & Reig-Martínez, E. (2011). Assessing farming eco-efficiency: A Data Envelopment Analysis approach. *Journal of Environmental Management*, 92(4), 1154–1164.

Picazo-Tadeo, A. J. & Reig-Martínez, E. (2006). Agricultural externalities and environmental regulation: Evaluating good practice in citrus production. *Applied Economics*, 38(11), 1327–1334.

Pielke, R. A. (2005). Land Use and Climate Change. *Science*, 310(5754), 1625–1626.

Pishgar-Komleh, S. H., Čechura, L., & Kuzmenko, E. (2021). Investigating the dynamic eco-efficiency in agriculture sector of the European Union countries. *Environmental Science and Pollution Research*, 28(35), 48942–48954.

Pishgar-Komleh, S. H., Zylowski, T., Rozakis, S., & Kozyra, J. (2020). Efficiency under different methods for incorporating undesirable outputs in an LCA+DEA framework: A case study of winter wheat production in Poland. *Journal of Environmental Management*, 260, 110138.

Pramova, E., Locatelli, B., Djoudi, H., & Somorin, O. A. (2012). Forests and trees for social adaptation to climate variability and change. *Wiley Interdisciplinary Reviews: Climate Change*, 3(6), 581–596.

Pröbstl-Haider, U., Mostegl, N. M., Kelemen-Finan, J., Haider, W., Formayer, H., Kantelhardt, J., Moser, T., Kapfer, M., & Trenholm, R. (2016). Farmers' Preferences for Future Agricultural Land Use Under the Consideration of Climate Change. *Environmental Management*, 58(3), 446–464.

Prokopy, L., Floress, K., Arbuckle, J., Church, S., Eanes, F., Gao, Y., Gramig, B., Ranjan, P., & Singh, A. (2019). Adoption of agricultural conservation practices in the United States: Evidence from 35 years of quantitative literature. *Journal of Soil and Water Conservation*, 74(5), 520–534.

Pufahl, A. & Weiss, C. R. (2009). Evaluating the effects of farm programmes: Results from propensity score matching. *European Review of Agricultural Economics*, 36(1), 79–101.

Pyhälä, A., Fernández-Llamazares, Á., Lehvävirta, H., Byg, A., Ruiz-Mallén, I., Salpeteur, M., & Thornton, T. F. (2016). Global environmental change: Local perceptions, understandings, and explanations. *Ecology and Society*, 21(3).

Qaim, M. (2020). Role of New Plant Breeding Technologies for Food Security and Sustainable Agricultural Development. *Applied Economic Perspectives and Policy*, 42(2), 129–150.

Ramos, D. d. L., Bustamante, M. M. C., Silva, Felipe D da Silva E, & Carvalheiro, L. G. (2018). Crop fertilization affects pollination service provision - Common bean as a case study. *PloS one*, 13(11), e0204460.

Ramsey, S. M., Bergtold, J. S., & Heier Stamm, J. L. (2021). FieldLevel LandUse Adaptation to Local Weather Trends. *American Journal of Agricultural Economics*, 103(4), 1314–1341.

Rana, P. & Miller, D. C. (2019). Machine learning to analyze the social-ecological impacts of natural resource policy: Insights from community forest management in the Indian Himalaya. *Environmental Research Letters*, 14(2).

Reinhard, S., Lovell, C. A. K., & Thijssen, G. (2002). Analysis of Environmental Efficiency Variation. *American Journal of Agricultural Economics*, 84, 1054–1065.

Ribeiro, M. T., Singh, S., & Guestrin, C. (2016). Model-agnostic interpretability of machine learning. URL: http://arxiv.org/abs/1606.05386 (last accessed: 2020-12-1).

Robinson, P. M. (1988). Root-N-Consistent Semiparametric Regression. *Econometrica*, 56(4), 931–954.

Rockström, J., Steffen, W., K. Noone, Å. Persson, Chapin, F., E. F. Lambin, T. M. Lenton, M. Scheffer, C. Folke, H. J. Schellnhuber, B. Nykvist, C. A. de Wit, T. Hughes, S. van der Leeuw, H. Rodhe, S. Sörlin, P. K. Snyder, R. Costanza, U. Svedin, M. Falkenmark, L. Karlberg, R. W. Corell, V. J. Fabry, J. Hansen, B. Walker, D. Liverman, K. Richardson, P. Crutzen, & J. A. Foley (2009). A safe operation space for humanity. *Nature*, 461(September), 472–475.

Rödl, A. (2017). Short Rotation Coppice: Status and Prospects. In R.A. Meyers (ed.) (Ed.), *Encyclopedia of Sustainability Science and Technology* (pp. 71–88). New York: Springer Science+Business Media.

Romanova, O., Gold, M., & Hendrickson, M. (2021). Temporal aspects of agroforestry adoption: SARE case study. *Agroforestry Systems*, 96(3), 659–668.

Rosa-Schleich, J., Loos, J., Mußhoff, O., & Tscharntke, T. (2019). Ecological-economic trade-offs of Diversified Farming Systems – A review. *Ecological Economics*, 160(January 2018), 251–263.

Rosenqvist, H. & Dawson, M. (2005). Economics of willow growing in Northern Ireland. *Biomass and Bioenergy*, 28(1), 7–14.

Rosenzweig, C. & Tubiello, F. N. (2007). Adaptation and mitigation strategies in agriculture: An analysis of potential synergies. *Mitigation and Adaptation Strategies for Global Change*, 12(5), 855–873.

Rubin, D. B. (1974). Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of Educational Psychology*, 66(5), 688–701.

Rubin, D. B. (1977). Assignment to treatment group on the basis of a covariate. *Educational Research*, 2(1), 1–26.

Ruto, E. & Garrod, G. (2009). Investigating farmers' preferences for the design of agri-environment schemes: a choice experiment approach. *Journal of Environmental Planning and Management*, 52(5), 631–647.

Salazar-Espinoza, C., Jones, S., & Tarp, F. (2015). Weather shocks and cropland decisions in rural Mozambique. *Food Policy*, 53, 9–21.

Salhofer, K. & Feichtinger, P. (2020). Regional differences in the capitalisation of first and second pillar payments of the CAP into land rental prices. *European Review of Agricultural Economics*, 48(1), 8–41.

Sauer, J., Frohberg, K., & Hockmann, H. (2006). Stochastic efficiency measurement: the curse of theoretical consistency. *Journal of Applied Economics*, 9(1), 139–165.

Sauer, J. & Moreddu, C. (2020). Drivers of farm performance: Empirical country case studies. In *ECD Food, Agriculture and Fisheries Papers, No. 143* (pp. 1–69). Paris: OECD Publishing.

Sauer, J. & Wossink, A. (2013a). Marketed outputs and non-marketed ecosystem services: The evaluation of marginal costs. *European Review of Agricultural Economics*, 40(4), 573–603.

Sauer, J. & Wossink, A. (2013b). Marketed outputs and non-marketed ecosystem services: The evaluation of marginal costs. *European Review of Agricultural Economics*, 40(4), 573–603.

Saunois, M., Jackson, R. B., Bousquet, P., Poulter, B., & Canadell, J. G. (2016). The growing role of methane in anthropogenic climate change. *Environmental Research Letters*, 11(12), 120207.

Scanlon, B. R., Jolly, I., Sophocleous, M., & Zhang, L. (2007). Global impacts of conversions from natural to agricultural ecosystems on water resources: Quantity versus quality. *Water Resources Research*, 43(3), 1–18.

Schaafsma, M., Ferrini, S., & Turner, R. K. (2019). Assessing smallholder preferences for incentivised climate-smart agriculture using a discrete choice experiment. *Land Use Policy*, 88(Novemeber 2019), 104153.

Schlenker, W., Hanemann, W. M., & Fisher, A. C. (2005). American Economic Association Will U . S . Agriculture Really Benefit from Global Warming ? Accounting for Irrigation in the Hedonic Approach. *The American Economic Review*, 95(1), 395–406.

Schlenker, W. & Roberts, M. J. (2009). Nonlinear temperature effects indicate severe damages to U.S. crop yields under climate change. *Proceedings of the National Academy of Sciences of the United States of America*, 106(37), 15594–15598.

Schmidt, O., Padel, S., & Levidow, L. (2012). The Bio-Economy Concept and Knowledge Base in a Public Goods and Farmer Perspective. *Bio-based and Applied Economics*, 1(1), 47–63.

Schoeneberger, M., Bentrup, G., De Gooijer, H., Soolanayakanahally, R., Sauer, T., Brandle, J., Zhou, X., & Current, D. (2012). Branching out: Agroforestry as a climate change mitigation and adaptation tool for agriculture. *Journal of Soil and Water Conservation*, 67(5).

Schoeneberger, M. M., Bentrup, G., & Patel-Weynand, T. (2017). Agroforestry : enhancing resiliency in U . S . agricultural landscapes under changing conditions. *United states Department of Agriculture*, (November 2017), 1–228.

Schomers, S. & Matzdorf, B. (2013). Payments for ecosystem services: A review and comparison of developing and industrialized countries. *Ecosystem Services*, 6, 16–30.

Schroeder, P. (1993). Agroforestry systems: integrated land use to store and conserve carbon. *Climate Research*, 3(1-2), 53–60.

Seibold, S., Gossner, M. M., Simons, N. K., Blüthgen, N., Müller, J., Ambarl\i, D., Ammer, C., Bauhus, J., Fischer, M., Habel, J. C., Linsenmair, K. E., Nauss, T., Penone, C., Prati, D., Schall, P., Schulze, E.-D., Vogt, J., Wöllauer, S., & Weisser, W. W. (2019). Arthropod decline in grasslands and forests is associated with landscape-level drivers. *Nature*, 574(7780), 671–674.

Serra, T., Chambers, R. G., & Oude Lansink, A. (2014). Measuring technical and environmental efficiency in a state-contingent technology. *European Journal of Operational Research*, 236(2), 706–717.

Serra, T., Lansink, A. O., & Stefanou, S. E. (2011). Measurement of dynamic efficiency: A directional distance function parametric approach. *American Journal of Agricultural Economics*, 93(3), 752–763.

Sexton, J. & Laake, P. (2009). Standard errors for bagged and random forest estimators. *Computational Statistics and Data Analysis*, 53(3), 801–811.

Shapley, L. S. (1988). The Shapley Value. In A. E. Roth (Ed.), *Contributions to the Theory of Games 2.28*, Annals of Mathematics Studies. Princeton: Princeton University Press, 2 edition.

Shepherd, R. W. (1970). *Theory of Cost and Production Functions*. Princeton: Princeton University Press.

Shortall, O. K. & Barnes, A. P. (2013). Greenhouse gas emissions and the technical efficiency of dairy farmers. *Ecological Indicators*, 29, 478–488.

Sinclair, K., Curtis, A., Mendham, E., & Mitchell, M. (2014). Can resilience thinking provide useful insights for those examining efforts to transform contemporary agriculture? *Agriculture and Human Values*, 31(3), 371–384.

Skevas, I., Emvalomatis, G., & Brümmer, B. (2018a). Heterogeneity of Long-run Technical Efficiency of German Dairy Farms: A Bayesian Approach. *Journal of Agricultural Economics*, 69(1), 58–75.

Skevas, I., Emvalomatis, G., & Brümmer, B. (2018b). Productivity growth measurement and decomposition under a dynamic inefficiency specification: The case of German dairy farms. *European Journal of Operational Research*, 271(1), 250–261.

Skevas, I., Emvalomatis, G., & Brümmer, B. (2018c). The effect of farm characteristics on the persistence of technical inefficiency: A case study in German dairy farming. *European Review of Agricultural Economics*, 45(1), 3–25.

Skevas, I., Zhu, X., Shestalova, V., & Emvalomatis, G. (2018d). The impact of agri-environmental policies and production intensification on the environmental performance of Dutch dairy farms. *Journal of Agricultural and Resource Economics*, 43(3), 423–440.

Smit, B. & Wandel, J. (2006). Adaptation, adaptive capacity and vulnerability. *Global Environmental Change*, 16(3), 282–292.

Smith, P., Goulding, K. W., Smith, K. A., Powlson, D. S., Smith, J. U., Falloon, P., & Coleman, K. (2001). Enhancing the carbon sink in European agricultural soils: Including trace gas fluxes in estimates of carbon mitigation potential. *Nutrient Cycling in Agroecosystems*, 60(1-3), 237–252.

Smith, P., Martino, D., Cai, Z., Gwary, D., Janzen, H., Kumar, P., McCarl, B., Ogle, S., O'Mara, F., Rice, C., Scholes, B., Sirotenko, O., Howden, M., McAllister, T., Pan, G., Romanenkov, V., Schneider, U., Towprayoon, S., Wattenbach, M., & Smith, J. (2008). Greenhouse gas mitigation in agriculture. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 363(1492), 789–813.

Smith, P. & Olesen, J. E. (2010). Synergies between the mitigation of, and adaptation to, climate change in agriculture. *Journal of Agricultural Science*, 148(5), 543–552.

Smukler, S. M., Sánchez-Moreno, S., Fonte, S. J., Ferris, H., Klonsky, K., O'Geen, A. T., Scow, K. M., Steenwerth, K. L., & Jackson, L. E. (2010). Biodiversity and multiple ecosystem functions in an organic farmscape. *Agriculture, Ecosystems and Environment*, 139(1-2), 80–97.

Soliman, T. & Djanibekov, U. (2021). Assessing dairy farming eco-efficiency in New Zealand: a two-stage data envelopment analysis. *New Zealand Journal of Agricultural Research*, 64(3), 411–428.

Sollen-Norrlin, M., Ghaley, B. B., & Rintoul, N. L. J. (2020). Agroforestry Benefits and Challenges for Adoption in Europe and Beyond. *Sustainability*, 12(17), 7001.

Song, J. & Chen, X. (2019). Eco-efficiency of grain production in China based on water footprints: A stochastic frontier approach. *Journal of Cleaner Production*, 236, 117685.

Song, W., Han, Z., & Deng, X. (2016). Changes in productivity, efficiency and technology of China's crop production under rural restructuring. *Journal of Rural Studies*, 47, 563–576.

Soteriades, A. D., Faverdin, P., Moreau, S., Charroin, T., Blanchard, M., & Stott, A. W. (2016). An approach to holistically assess (dairy) farm eco-efficiency by combining Life Cycle Analysis with Data Envelopment Analysis models and methodologies. *Animal*, 10(11), 1899–1910.

Soteriades, A. D., Foskolos, A., Styles, D., & Gibbons, J. M. (2020). Maintaining production while reducing local and global environmental emissions in dairy farming. *Journal of Environmental Management*, 272(July), 111054.

Spiegel, A., Soriano, B., de Mey, Y., Slijper, T., Urquhart, J., Bardají, I., Vigani, M., Severini, S., & Meuwissen, M. (2020). Risk Management and its Role in Enhancing Perceived Resilience Capacities of Farms and Farming Systems in Europe. *EuroChoices*, 19(2), 45–53.

Staniszewski, J. (2018). Attempting to measure sustainable intensification of agriculture in countries of the European Union. *Journal of Environmental Protection and Ecology*, 19(2), 949–957.

Stefanou, S. E. (2009). A Dynamic Characterization of Efficiency. *Agricultural economics review*, 10(1), 18–33.

Stępień, S., Czyżewski, B., Sapa, A., Borychowski, M., Poczta, W., & Poczta-Wajda, A. (2021). Eco-efficiency of small-scale farming in Poland and its institutional drivers. *Journal of Cleaner Production*, 279, 123721.

Stetter, C., Mennig, P., & Sauer, J. (2022a). Using Machine Learning to Identify Heterogeneous Impacts of Agri-Environment Schemes in the EU: A Case Study. *European Review of Agricultural Economics*.

Stetter, C. & Sauer, J. (2021). Exploring the heterogeneous effects of weather on productivity using generalized random forests. In *95th Annual Conference of the Agricultural Economics Society*.

Stetter, C. & Sauer, J. (2022). Greenhouse Gas Emissions and Eco-Performance at Farm Level: A Parametric Approach. *Environmental and Resource Economics*, 81(3), 617–647.

Stetter, C., Wimmer, S., & Sauer, J. (2022b). Are Intensive Farms More Emission-Efficient? Evidence From German Dairy Farms. *Journal of Agricultural and Resource Economics*.

StMELF (2018). *Bayerisches Kulturlandschaftsprogramm (KULAP) und Bayerisches Vertragsnaturschutzprogramm inkl. Erschwernisausgleich (VNP): Merkblatt 2019 bis 2023 Agrarumwelt- und Klimamaßnahmen (AUM)*. URL: https://www.stmelf.bayern.de/mam/cms01/agrarpolitik/dateien/m_aum_-verpflichtungszeitraum_2019_2023.pdf (last accessed: 2022-6-7). Munich: Bavarian Ministry of Food, Agriculture and Forestry.

StMELF (2021). *Bayerischer Agrarbericht 2020*. URL: https://www.agrarbericht.bayern.de/politik-strategien/index.html (last accessed: 2022-6-7). Munich: Bavarian Ministry of Food, Agriculture and Forestry.

Storm, H., Baylis, K., & Heckelei, T. (2020). Machine learning in agricultural and applied economics. *European Review of Agricultural Economics*, 47(3), 849–892.

Štrumbelj, E. & Kononenko, I. (2014). Explaining prediction models and individual predictions with feature contributions. *Knowledge and Information Systems*, 41(3), 647–665.

Styles, D., Gonzalez-Mejia, A., Moorby, J., Foskolos, A., & Gibbons, J. (2018). Climate mitigation by dairy intensification depends on intensive use of spared grassland. *Global change biology*, 24(2), 681–693.

Tanner, E. P. & Fuhlendorf, S. D. (2018). Impact of an agri-environmental scheme on landscape patterns. *Ecological Indicators*, 85(November 2017), 956–965.

Tiffin, A. (2019). *Machine learning and causality: The impact of financial crises on growth*. IMF Working Papers. WP/19/228. Washington: International Monetary Fund.

Tilman, D., Cassman, K. G., Matson, P. A., Naylor, R., & Polasky, S. (2002). Agricultural sustainability and intensive production practices. *Nature*, 418(6898), 671–677.

Tilman, D., Fargione, J., Wolff, B., D'Antonio, C., Dobson, A., Howarth, R., Schindler, D., Schlesinger, W. H., Simberloff, D., & Swackhamer, D. (2001). Forecasting agriculturally driven global environmental change. *Science*, 292(5515), 281–284.

Tilman, D. & Lehman, C. (2001). Human-caused environmental change: Impacts on plant diversity and evolution. *Proceedings of the National Academy of Sciences of the United States of America*, 98(10), 5433–5440.

Tinbergen, J. (1956). *Economic Policy: Principles and Design*. Erasmus School of Economics. URL: http://hdl.handle.net/1765/16740 (last accessed: 2021-06-09). Amsterdam: North-Holland.

Tomich, T. P., Brodt, S., Ferris, H., Galt, R., Horwath, W. R., Kebreab, E., Leveau, J. H., Liptzin, D., Lubell, M., Merel, P., Michelmore, R., Rosenstock, T., Scow, K., Six, J., Williams, N., & Yang, L. (2011). Agroecology: A review from a global-change perspective. *Annual Review of Environment and Resources*, 36(1), 193–222.

Train, K. E. (2001). *Discrete Choice Methods with Simulation*. New York: Cambridge University Press.

Tsagris, M. & Tzouvelekas, V. (2022). Nitrate leaching and efficiency measurement in intensive farming systems: A parametric byproduction technology approach. *Agricultural Economics*, (November), 1–15.

Tsiafouli, M. A., Thébault, E., Sgardelis, S. P., de Ruiter, P. C., van der Putten, W. H., Birkhofer, K., Hemerik, L., de Vries, F. T., Bardgett, R. D., Brady, M. V., Bjornlund, L., Jørgensen, H. B., Christensen, S., Hertefeldt, T. D., Hotes, S., Gera Hol, W. H., Frouz, J., Liiri, M., Mortimer, S. R., Setälä, H., Tzanopoulos, J., Uteseny, K., Pižl, V., Stary, J., Wolters, V., & Hedlund, K. (2015). Intensive agriculture reduces soil biodiversity across Europe. *Global Change Biology*, 21(2), 973–985.

Tsionas, E. G. (2002). Stochastic frontier models with random coefficients. *Journal of Applied Econometrics*, 17(2), 127–147.

Tsionas, E. G. (2006). Inference in dynamic stochastic frontier models. *Journal of Applied Econometrics*, 21(5), 669–676.

Tsonkova, P., Böhm, C., Quinkenstein, A., & Freese, D. (2012). Ecological benefits provided by alley cropping systems for production of woody biomass in the temperate region: A review. *Agroforestry Systems*, 85(1), 133–152.

Tyteca, D. (1996). On the Measurement of the Environmental Performance of Firms - A Literature Review and a Productive Efficiency Perspective. *Journal of Environmental Management*, 46(3), 281–308.

Tyteca, D. (1998). Sustainability Indicators at the Firm Level Pollution and Resource Efficiency as a Necessary Condition Toward Sustainability. *Journal of Industrial Ecology*.

Uehleke, R., Petrick, M., & Hüttel, S. (2022). Evaluations of agri-environmental schemes based on observational farm data: The importance of covariate selection. *Land Use Policy*, 114, 105950.

Ullah, A., Perret, S. R., Gheewala, S. H., & Soni, P. (2016). Eco-efficiency of cotton-cropping systems in Pakistan: an integrated approach of life cycle assessment and data envelopment analysis. *Journal of Cleaner Production*, 134(Part B), 623–632.

UNFCCC (2014). *Guidelines for the preparation of national communications by Parties included in Annex I to the Convention, Part I: UNFCCC reporting guidelines on annual greenhouse gas inventories.* United Nations FCCC/CP/2013/10/Add.3. Decision 24/CP.19 Revision. Warsaw: United Nations Framework Convention on Climate Change.

USDA (2019). *Agroforestry Strategic Framework. Fiscal Years 2019–2024.* Washington: U.S. Department of Agriculture (USDA).

USDA (2022). TFP indices and components for countries, regions, countries grouped by income level, and the world, 1961-2019. URL: www.ers.usda.gov/data-products/international-agricultural-productivity/ (last accessed: 2022-04-27).

Uthes, S., Matzdorf, B., Müller, K., & Kaechele, H. (2010). Spatial targeting of agri-environmental measures: cost-effectiveness and distributional consequences. *Environmental management*, 46(3), 494–509.

Valentine, J., Clifton-Brown, J., Hastings, A., Robson, P., Allison, G., & Smith, P. (2012). Food vs. fuel: The use of land for lignocellulosic 'next generation' energy crops that minimize competition with primary food production. *GCB Bioenergy*, 4(1), 1–19.

van der Horst, D. (2007). Assessing the efficiency gains of improved spatial targeting of policy interventions; the example of an agri-environmental scheme. *Journal of Environmental Management*, 85(4), 1076–1087.

van Dijk, M., Morley, T., Rau, M. L., & Saghai, Y. (2021). A meta-analysis of projected global food demand and population at risk of hunger for the period 2010–2050. *Nature Food*, 2(7), 494–501.

van Noordwijk, M., Bayala, J., Hairiah, K., Lusiana, B., Muthuri, C., Khasanah, N., & Mulia, R. (2014). Agroforestry solutions for buffering climate variability and adapting to change. In *Climate change impact and adaptation in agricultural systems* (pp. 216–232). Wallingford: CABI.

van Noordwijk, M., Coe, R., Sinclair, F. L., Luedeling, E., Bayala, J., Muthuri, C. W., Cooper, P., Kindt, R., Duguma, L., Lamanna, C., & Minang, P. A. (2021). Climate change adaptation in and through agroforestry: four decades of research initiated by Peter Huxley. *Mitigation and Adaptation Strategies for Global Change*, 26(5), 18.

van Noordwijk, M., Hoang, M., Neufeldt, H., Öborn, I., & Yataich, T., Eds. (2011). *How trees and people can co-adapt to climate change: reducing vulnerability through multifunctional agroforestry landscapes.* Nairobi: World Agroforestry Centre (ICRAF).

Verchot, L. V., Van Noordwijk, M., Kandji, S., Tomich, T., Ong, C., Albrecht, A., Mackensen, J., Bantilan, C., Anupama, K. V., & Palm, C. (2007). Climate change: Linking adaptation and mitigation through agroforestry. *Mitigation and Adaptation Strategies for Global Change*, 12(5), 901–918.

Vergamini, D., Viaggi, D., & Raggi, M. (2020). Evaluating the Potential Contribution of Multi-Attribute Auctions to Achieve Agri-Environmental Targets and Efficient Payment Design. *Ecological Economics*, 176(June), 106756.

Viney, R., Savage, E., & Louviere, J. (2005). Empirical investigation of experimental design properties of discrete choice experiments in health care. *Health Economics*, 14(4), 349–362.

Vitousek, P. M. (1992). Global Environmental Change: An Introduction. *Annual Review of Ecology and Systematics*, 23(1), 1–14.

Vogel, E., Donat, M. G., Alexander, L. V., Meinshausen, M., Ray, D. K., Karoly, D., Meinshausen, N., & Frieler, K. (2019). The effects of climate extremes on global agricultural yields. *Environmental Research Letters*, 14(5), 054010.

Vroege, W., Vrieling, A., & Finger, R. (2021). Satellite support to insure farmers against extreme droughts. *Nature Food*, 1-3.

Vrolijk, H. & Poppe, K. (2021). Cost of Extending the Farm Accountancy Data Network to the Farm Sustainability Data Network: Empirical Evidence. *Sustainability*, 13(15), 8181.

Wager, S. & Athey, S. (2018). Estimation and Inference of Heterogeneous Treatment Effects using Random Forests. *Journal of the American Statistical Association*, 113(523), 1228–1242.

Wager, S., Hastie, T., & Efron, B. (2013). Confidence Intervals for Random Forests: The Jackknife and the Infinitesimal Jackknife. *Journal of Machine Learning Research*, 15, 1625–1651.

Wang, H. J. & Ho, C. W. (2010). Estimating fixed-effect panel stochastic frontier models by model transformation. *Journal of Econometrics*, 157(2), 286–296.

Wang, Y. & Blei, D. M. (2019). The blessings of multiple causes. *Journal of the American Statistical Association*, 114(528), 1574–1596.

Wätzold, F., Drechsler, M., Johst, K., Mewes, M., & Sturm, A. (2016). A novel, spatiotemporally explicit ecologicaleconomic modeling procedure for the design of costeffective agrienvironment schemes to conserve biodiversity. *American Journal of Agricultural Economics*, 98(2), 489–512.

Webber, H., Lischeid, G., Sommer, M., Finger, R., Nendel, C., Gaiser, T., & Ewert, F. (2020). No perfect storm for crop yield failure in Germany. *Environmental Research Letters*, 15(10), 104012.

WEF (2020). *The Global Risks Report 2020*. World Economic Forum. URL: http://wef.ch/risks2019 (last accessed 2021-08-31).

Weiske, A., Vabitsch, A., Olesen, J. E., Schelde, K., Michel, J., Friedrich, R., & Kaltschmitt, M. (2006). Mitigation of greenhouse gas emissions in European conventional and organic dairy farming. *Agriculture, Ecosystems & Environment*, 112(2-3), 221–232.

Weitzman, M. L. (2012). GHG Targets as Insurance Against Catastrophic Climate Damages. *Journal ofPublic Economic Theory*, 14(2), 221–244.

Westbury, D. B., Park, J. R., Mauchline, A. L., Crane, R. T., & Mortimer, S. R. (2011). Assessing the environmental performance of English arable and livestock holdings using data from the Farm Accountancy Data Network (FADN). *Journal of Environmental Management*, 92(3), 902–909.

Westerink, J., Jongeneel, R., Polman, N., Prager, K., Franks, J., Dupraz, P., & Mettepenningen, E. (2017). Collaborative governance arrangements to deliver spatially coordinated agri-environmental management. *Land Use Policy*, 69, 176–192.

Westerink, J., Melman, D. C. P., & Schrijver, R. A. M. (2014). Scale and self-governance in agri-environment schemes: Experiences with two alternative approaches in the Netherlands. *Journal of Environmental Planning and Management*, 58(8), 1490–1508.

Westra, S., Fowler, H. J., Evans, J. P., Alexander, L. V., Berg, P., Johnson, F., Kendon, E. J., Lenderink, G., & Roberts, N. M. (2014). Future changes to the intensity and frequency of short-duration extreme rainfall. *Reviews of Geophysics*, 52(3), 522–555.

Wilson, R. S., Herziger, A., Hamilton, M., & Brooks, J. S. (2020). From incremental to transformative adaptation in individual responses to climate-exacerbated hazards. *Nature Climate Change*, 10(3), 200–208.

Wimmer, S. & Sauer, J. (2020). Diversification economies in dairy farming – empirical evidence from Germany. *European Review of Agriculture Economics*, 47(3), 1338–1368.

Winkler, K., Fuchs, R., Rounsevell, M., & Herold, M. (2021). Global land use changes are four times greater than previously estimated. *Nature Communications*, 12(1), 1–10.

Wolbert-Haverkamp, M. & Musshoff, O. (2014). Are short rotation coppices an economically interesting form of land use? A real options analysis. *Land Use Policy*, 38, 163–174.

Wolz, K. J., Lovell, S. T., Branham, B. E., Eddy, W. C., Keeley, K., Revord, R. S., Wander, M. M., Yang, W. H., & DeLucia, E. H. (2018). Frontiers in alley cropping: Transformative solutions for temperate agriculture. *Global Change Biology*, 24(3), 883–894.

Wooldridge, J. M. (2005). Violating ignorability of treatment by controlling for too many factors. *Econometric Theory*, 21(05), 1026–1028.

Wossink, A. & Swinton, S. M. (2007). Jointness in production and farmers' willingness to supply non-marketed ecosystem services. *Ecological Economics*, 64(2), 297–304.

Wuepper, D. & Huber, R. (2021). Comparing effectiveness and return on investment of action and resultsbased agrienvironmental payments in <scp>S</scp> witzerland. *American Journal of Agricultural Economics*, (pp. 1–20).

Wuepper, D., Wimmer, S., & Sauer, J. (2020). Is small family farming more environmentally sustainable? Evidence from a spatial regression discontinuity design in Germany. *Land Use Policy*, 90(November), 104360.

Wunder, S., Börner, J., Ezzine-De-Blas, D., Feder, S., & Pagiola, S. (2020). Payments for environmental services: Past performance and pending potentials. *Annual Review of Resource Economics*, 12, 209–234.

Wünscher, T., Engel, S., & Wunder, S. (2008). Spatial targeting of payments for environmental services: A tool for boosting conservation benefits. *Ecological Economics*, 65(4), 822–833.

Yang, B., Wang, Z., Zou, L., Zou, L., & Zhang, H. (2021). Exploring the eco-efficiency of cultivated land utilization and its influencing factors in China's Yangtze River Economic Belt, 2001–2018. *Journal of Environmental Management*, 294(November 2020), 112939.

Zehetmeier, M., Hoffmann, H., Sauer, J., Hofmann, G., Dorfner, G., & O'Brien, D. (2014). A dominance analysis of greenhouse gas emissions, beef output and land use of German dairy farms. *Agricultural Systems*, 129, 55–67.

Zehetmeier, M., Läpple, D., Hoffmann, H., Zerhusen, B., Strobl, M., Meyer-Aurich, A., & Kapfer, M. (2020). Is there a joint lever? Identifying and ranking factors that determine GHG emissions and profitability on dairy farms in Bavaria, Germany. *Agricultural Systems*, 184, 102897.

Zhang, W., Ricketts, T. H., Kremen, C., Carney, K., & Swinton, S. M. (2007). Ecosystem services and dis-services to agriculture. *Ecological Economics*, 64(2), 253–260.

Zilberman, D., Zhao, J., & Heiman, A. (2012). Adoption Versus Adaptation, with Emphasis on Climate Change. *Annual Review of Resource Economics*, 4(1), 27–53.

Zimmermann, A. & Britz, W. (2016). European farms' participation in agri-environmental measures. *Land Use Policy*, 50, 214–228.

Supplementary material

Author agreement with *Environmental and Resource Economics* containing the authorization to include the manuscript in this thesis.

Licence to Publish - Open Access

Licensee:	Springer Nature B.V.	(the 'Licensee')
Journal Name:	Environmental and Resource Economics	(the 'Journal')
Manuscript Number:	EARE-D-20-00322R1	
Proposed Title of Article:	Greenhouse Gas Emissions and Eco-Performance at Farm Level: A Parametric Approach	(the 'Article')
Author(s) [Please list all named Authors]:	Christian Stetter, Johannes Sauer	(the 'Author')
Corresponding Author Name:	Christian Stetter	

Licence Applicable to the Article:

Creative Commons licence CC BY: This licence allows readers to copy, distribute and transmit the Article as long as it is attributed back to the author. Readers are permitted to alter, transform or build upon the Article, and to use the Article for commercial purposes. Please read the full licence for further details at - http://creativecommons.org/licenses/by/4.0/

Subject to editorial acceptance of the Article, it will be published under the Creative Commons licence shown above.

1 Grant of Rights

a) For good and valuable consideration, the Author hereby grants to the Licensee the perpetual, non-exclusive, irrevocable, world-wide, assignable, sublicensable and unlimited right to: publish, reproduce, copy, distribute, communicate, display publicly, sell, rent and/ or otherwise make available the article identified above, including any supplementary information and graphic elements therein (e.g. illustrations, charts, moving images) (the "Article") in any language, in any versions or editions in any and all forms and/or media of expression (including without limitation in connection with any and all end-user devices), whether now known or developed in the future. Without limitation, the above grant includes: (i) the right to edit, alter, adapt, adjust and prepare derivative works; (ii) all advertising and marketing rights including without limitation in relation to social media; (iii) rights for any training, educational and/or instructional purposes; (iv) the right to add and/or remove links or combinations with other media/works; and (v) the right to create, use and/or license and/or sublicense content data or metadata of any kind in relation to the Article (including abstracts and summaries) without restriction. The above rights are granted in relation to the Article as a whole or any part and with or in relation to any other works.

b) Without limiting the rights granted above, Licensee is granted the rights to use the Article for the purposes of analysis, testing, and development of publishing- and research-related workflows, systems, products, projects, and services; to confidentially share the Article with select third parties to do the same; and to retain and store the Article and any associated correspondence/files/forms to maintain the historical record, and to facilitate research integrity investigations. The grant of rights set forth in this clause (b) is irrevocable.

c) The Licensee will have the right, but not the obligation, to exercise any or all of the rights granted herein. If the Licensee elects not to publish the Article for any reason, all publishing rights under this Agreement as set forth in clause 1.a) above will revert to the Author.

2 Copyright

Ownership of copyright in the Article will be vested in the name of the Author. When reproducing the Article or extracts from it, the Author will acknowledge and reference first publication in the Journal.

3 Use of Article Versions

a) For purposes of this Agreement: (i) references to the "Article" include all versions of the Article; (ii) "Submitted Manuscript" means the version of the Article as first submitted by the Author; (iii) "Accepted Manuscript" means the version of the Article accepted for publication, but prior to copy-editing and typesetting; and (iv) "Version of Record" means the version of the Article published by the Licensee, after copy-editing and typesetting. Rights to all versions of the Manuscript are granted on a non-exclusive basis.

b) The Author may make the Submitted Manuscript available at any time and under any terms (including, but not limited to, under a CC BY licence), at the Author's discretion. Once the Article has been published, the Author will include an acknowledgement and provide a link to the Version of Record on the publisher's website: "This preprint has not undergone peer review (when applicable) or any post-submission improvements or corrections. The Version of Record of this article is published in [insert journal title], and is available online at https://doi.org/[insert DOI]".

c) Immediately after acceptance the Author may deposit the Accepted Manuscript to any location, and under any terms (including, but not limited to, under a CC BY licence), provided it is not made publicly available until after publication. The Author will include an acknowledgement in the Accepted Manuscript, together with a link to the Version of Record on the publisher's website: "This version of the article has been accepted for publication, after peer review (when applicable) but is not the Version of Record and does not reflect post-acceptance improvements, or any corrections. The Version of Record is available online at: http://dx.doi.org/[insert DOI]".

4 Warranties & Representations

Author warrants and represents that:

a)

i. the Author is the sole copyright owner or has been authorised by any additional copyright owner(s) to grant the rights defined in clause 1,

ii. the Article does not infringe any intellectual property rights (including without limitation copyright, database rights or trade mark rights) or other third party rights and no licence from or payments to a third party are required to publish the Article,

iii. the Article has not been previously published or licensed, nor has the Author committed to licensing any version of the Article under a licence inconsistent with the terms of this Agreement,

iv. if the Article contains materials from other sources (e.g. illustrations, tables, text quotations), Author has obtained written permissions to the extent necessary from the copyright holder(s), to license to the Licensee the same rights as set out in clause 1 and has cited any such materials correctly;

b) all of the facts contained in the Article are according to the current body of research true and accurate;

c) nothing in the Article is obscene, defamatory, violates any right of privacy or publicity, infringes any other human, personal or other rights of any person or entity or is otherwise unlawful and that informed consent to publish has been obtained for any research participants;

d) nothing in the Article infringes any duty of confidentiality owed to any third party or violates any contract, express or implied, of the Author;

e) all institutional, governmental, and/or other approvals which may be required in connection with the research reflected in the Article have been obtained and continue in effect;

f) all statements and declarations made by the Author in connection with the Article are true and correct; and

g) the signatory who has signed this agreement has full right, power and authority to enter into this agreement on behalf of all of the Authors.

5 Cooperation

a) The Author will cooperate fully with the Licensee in relation to any legal action that might arise from the publication of the Article, and the Author will give the Licensee access at reasonable times to any relevant accounts, documents and records within the power or control of the Author. The Author agrees that any Licensee affiliate through which the Licensee exercises any rights or performs any obligations under this Agreement is intended to have the benefit of and will have the right to enforce the terms of this Agreement.

b) Author authorises the Licensee to take such steps as it considers necessary at its own expense in the Author's name(s) and on their behalf if the Licensee believes that a third party is infringing or is likely to infringe copyright in the Article including but not limited to initiating legal proceedings.

6 Author List

Changes of authorship, including, but not limited to, changes in the corresponding author or the sequence of authors, are not permitted after acceptance of a manuscript.

7 Post Publication Actions

The Author agrees that the Licensee may remove or retract the Article or publish a correction or other notice in relation to the Article if the Licensee determines that such actions are appropriate from an editorial, research integrity, or legal perspective.

8 Controlling Terms

The terms of this Agreement will supersede any other terms that the Author or any third party may assert apply to any version of the Article.

9 Governing Law

This Agreement will be governed by, and construed in accordance with, the laws of The Netherlands. The courts of Amsterdam, The Netherlands will have exclusive jurisdiction.

Springer Nature B.V., Van Godewijckstraat 30, 3311 GX Dordrecht, The Netherlands
v.3.1.3 - (09_2021)-SNBV

Author agreement with the *Journal of Agricultural and Resource Economics* containing the authorization to include the manuscript in this thesis.

JOURNAL OF AGRICULTURAL AND RESOURCE ECONOMICS

Editors:
Dr. Dragan Miljkovic, dragan.miljkovic@ndsu.edu
Dr. Anton Bekkerman, anton.bekkerman@montana.edu
Dr. Richard Woodward, r-woodward@tamu.edu
Dr. Darren Hudson, darren.hudson@ttu.edu

Publications Editor:
Amy Bekkerman
amy@precisionedits.com

AUTHOR(S) STATEMENT

For Manuscript Entitled:

Are Intensive Farms More Emission-Efficient? Evidence From German Dairy Farms

1. To the best of my knowledge, the material in this manuscript does not infringe upon other published material protected by copyright.
2. I grant the Western Agricultural Economics Association and the *Journal of Agricultural and Resource Economics* (hereafter *JARE*) the right of first publication and a nonexclusive license to publish and distribute the above-named article under a Creative Commons Attribution-NonCommercial 4.0 International License
3. *JARE* in turn acknowledges that the author(s) retain(s) copyright and publishing rights without restrictions. Authors may enter into separate, additional contractual arrangements for the non-exclusive distribution of *JARE*'s published version of the work, including but not limited to reproducing the contribution separately, in a university or institutional repository, or in any derivative work of which they are the author(s) or editor(s), subject only to giving proper credit in the work to the original publication of the contribution by *JARE*.
4. I understand that I will receive no monetary return from *JARE* for the use of the material.
5. I agree to report, describe, and/or reference the complete estimation process undertaken to derive the results presented in the manuscript, including data documentation as well as model specifications estimated but not reported in the manuscript. I agree to make data used for this manuscript available at cost upon request for five years.
6. I agree to include a footnote disclosing any and all funding sources that supported the research described in this manuscript.
7. I agree to assume obligation for payment of page charges at the time of publication. Current page charges are $79/printed page for WAEA members and $89/printed page for nonmembers, plus $10/page for online supplementary appendices.

Author #1: [redacted] Date November 9, 2021

Author #2: [redacted] Date November 9, 2021

Author #3: [redacted] Date November 10, 2021

Author #4: ______ Date ______

☐ Check this box if one or more of the authors is an employee of the U.S. Government and the referenced article was written in that capacity. *JARE* acknowledges that this statement applies only to the extent allowable by U.S. Law.

☑ Check this box if the data used in the referenced article are proprietary and unavailable to release for reanalysis. *JARE* acknowledges that the author(s) cannot make data used for this manuscript available in this case.

Author agreement with the *European Review of Agricultural Economics* containing the authorization to include the manuscript in this thesis.

Journal: European Review of Agricultural Economics
DOI: 10.1093/erae/jbab057
Title: Using Machine Learning to Identify Heterogeneous Impacts of Agri-Environment Schemes in the EU: A Case Study

Standard Licence

You hereby grant to Oxford University Press and The Foundation for the European Review of Agricultural Economics an exclusive licence for the full period of copyright throughout the world:

- to publish the final version of the Article in the above Journal, and to distribute it and/or to communicate it to the public, either within the Journal, on its own, or with other related material throughout the world, in printed, electronic or any other format or medium whether now known or hereafter devised;
- to make translations and abstracts of the Article and to distribute them to the public;
- to authorize or grant licences to third parties to do any of the above;
- to deposit copies of the Article in online archives maintained by OUP or by third parties authorized by OUP.

Â

You authorize us to act on your behalf to defend the copyright in the Article if anyone should infringe it and to register the copyright of the Article in the US and other countries, if necessary.In the case of a multi authored article, you confirm that you are authorized by your co-authors to enter the licence on their behalf.

Â

You confirm to OUP that the Article:

- is your original work;
- has not previously been published (in print or electronic format), is not currently under consideration by another journal, or if it has already been submitted to other journal, it will be immediately withdrawn;
- will not be submitted for publication to any other journal following acceptance in the above Journal; and
- OUP will be the first publisher of the Article.

Â

You warrant to OUP that:

- no part of the Article is copied from any other work,
- you have obtained ALL the permissions required (for print and electronic use) for any material you have used from other copyrighted publications in the Article; and
- you have exercised reasonable care to ensure that the Article is accurate and does not contain anything which is libellous, or obscene, or infringes on anyone's copyright, right of privacy, or other rights.

Â

RETAINED RIGHTS

Abstract and Citation informationÂ - Authors may reuse the Abstract and Citation information (e.g. Title, Author name, Publication dates) of their article anywhere at any time including social media such as Facebook, blogs and Twitter, providing that where possible a link is included back to the article on the OUP site. Preferably the link should be, or include, the Digital Object Identifier (DOI) which can be found in the Citation information about your article online.Â

Authorâ s Original VersionÂ - Authors may reuse the Authorâ s Original Version (AOV) anywhere at any time, providing that once the article is accepted they provide a statement of acknowledgement, and that once the article has been published this acknowledgement is updated to provide details such as the volume and issue number, the DOI, and a link to the published article on the journalâ s websiteÂ

Accepted ManuscriptÂ - On first online publication authors may immediately upload their Accepted Manuscript (AM) to their own personal webpage, or their institutional repository on the proviso that it is not made publically available untilÂ **24Â months**Â after the online date of publication. After the embargo period, authors may make their AM publicly available through their institutional repository or other non-commercial repositories PROVIDED THAT (1) you include a link (URL) to the VERSION OF RECORD on the journalsâ website; and (2) the Journal is attributed as the original place of publication with the correct citation details given. AM may not be uploaded to commercial websites or repositories, unless the website or repository has signed a licensing agreement with OUP.Â Â

Version of RecordÂ - The Version of Record (VOR) as it appears in the journal following copyediting and proof correction may not be deposited by authors in institutional repositories or posted to third party websites unlessÂ the repository or website has signed an agreement with OUP allowing posting.Â

Free link to published articleÂ - On publication, authors will be sent an online offprint link allowing access to their article on the OUP website without subscription. For authors of non- Open Access articles, this link may be shared directly with interested colleagues, but is not intended for mass distribution through, repositories, or social media. If you wish to share links or publicize your article we would ask that you instead distribute a link to the abstract of the article.Â

Funding bodiesÂ - Please be aware that you are responsible for all funding agency compliance and the accuracy of information you provide in relation to your article. OUP and/or the controlling Learned Society shall not be responsible for checking that funding agency requirements have been complied with.Â

Educational useÂ - You may use the Article within your employerâ s institution or company for educational or research purposes only, including use in course-packs, as long as: (1) you do not use it for commercial purposes or re-distribution outside of the institution/company; (2) you acknowledge the Journal as the original place of publication with the correct citation details given.Â

Further informationÂ - details of OUPâ s publication rights policies, including author rights can be found atÂ https://academic.oup.com/journals/pages/access_purchase/rights_and_permissions/publication_rights

Â

christian.stetter@tum.de signed this licence on 2021-12-28 08:36:01 GMT.