



HAL
open science

Modelling farmers' production choices and chemical inputs demand with a latent function approach

Esther Devilliers

► **To cite this version:**

Esther Devilliers. Modelling farmers' production choices and chemical inputs demand with a latent function approach. Economics and Finance. Institut national d'enseignement supérieur pour l'agriculture, l'alimentation et l'environnement, 2021. English. ⟨NNT : 2021AGROE058⟩. ⟨tel-03555413v2⟩

HAL Id: tel-03555413

<https://hal.science/tel-03555413v2>

Submitted on 3 Feb 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization

THESE DE DOCTORAT DE

*L'institut national d'enseignement supérieur pour l'agriculture, l'alimentation
et l'environnement*
Ecole interne AGROCAMPUS OUEST

ECOLE DOCTORALE N° 597
Sciences Economiques et sciences De Gestion
Spécialité : « *Sciences économiques* »

Par

Esther DEVILLIERS

**Modélisation micro-économétrique des choix de pratiques de
production et des utilisations d'intrants chimiques des agriculteurs :
une approche par les fonctions de production latentes**

Thèse présentée et soutenue à Rennes, le 3 septembre 2021
Unité de recherche : UMR SMART-LERECO
Thèse N° : E58_2021-18

Rapporteurs avant soutenance :

Stéphane Auray Professeur en Économie, ENSAI
Jean-Paul Chavas Professor in Agricultural and Applied Economics, University of Wisconsin-Madison

Composition du Jury :

Président : Stéphane Auray Professeur en Économie, ENSAI
Examineurs : Sabine Duvaléix Maître de Conférences en Économie Agricole, Agrocampus Ouest
 Robert Finger Professor of Agricultural Economics and Policy, ETH Zürich
Dir. de thèse : Alain Carpentier Directeur de Recherche en Économie Agricole, INRAE

Invité(s)

This Ph.D thesis should not be reported as representing the views of l'Institut Agro - AGROCAMPUS OUEST. The views expressed are those of the author and do not necessarily reflect those of the University.

L'institut Agro - AGROCAMPUS OUEST n'entend donner aucune approbation ni improbation aux opinions émises dans cette thèse. Ces opinions doivent être considérées comme propres à leur auteure.

"At the end of the day, we can endure much more than we think we can."

Frida Kahlo

"We choose to go to the moon in this decade and do the other things, not because they are easy, but because they are hard, because that goal will serve to organize and measure the best of our energies and skills, because that challenge is one that we are willing to accept, one we are unwilling to postpone, and one we intend to win."

John F. Kennedy

Remerciements

Un manuscrit de thèse, ce n'est pas seulement faire le bilan de trois (quatre pour ma part) années de recherche... Un manuscrit de thèse, c'est mettre un point final à trois années de vie, à trois années de cheminement tant d'un point de vue intellectuel que personnel. Pour ce qui est du cheminement intellectuel, vous aurez l'occasion d'en juger par vous-même à la lecture des 150 pages qui constituent mon manuscrit. Pour ce qui est du cheminement personnel, je voudrais prendre ici quelques lignes pour remercier les personnes qui m'ont accompagnées durant ces trois années.

À celui d'abord qui a osé croire en moi en première instance, mon directeur de thèse, Alain Carpentier. J'en profite également pour remercier ceux qui ont financé ce travail de thèse, à savoir le département SAE2 de l'INRA et le projet ANR Soilserv. Sans vous, ce travail n'aurait su voir le jour.

À ceux qui m'ont fait confiance par la suite. Robert Finger, en m'accueillant au sein de son groupe à l'ETH Zürich pour un séjour de recherche de 4 mois. L'ENSAI, en me permettant d'animer des TD en économie et en statistiques. Ronan Le Saout et l'ensemble du département d'économie de l'ENSAI, qui m'ont renouvelé leur confiance en me recrutant pour un poste d'ATER en économie.

À ceux également qui ont accepté de faire partie de mon jury de thèse. Je voudrais notamment remercier mes rapporteurs, Stéphane Auray et Jean-Paul Chavas, ainsi que mes examinateurs, Sabine Duvaléix, Salima El Kolei et Robert Finger.

À ceux qui ont accepté de me donner un peu de leur temps pour relire une partie de ce manuscrit, Enora Belz, Lucile Henry et Matthieu Marbac-Lourdelle. Merci pour vos précieux commentaires et retours.

À ceux qui sont d'ores-et-déjà, et certainement à jamais, associé à l'expérience de la thèse. Aux rencontres éphémères faites lors de conférences, lors d'écoles d'été ou lors de cours dispensés par l'école doctorale. Et aux rencontres plus pérennes...

Une pensée spéciale pour les anciens et actuels doctorants de l'unité SMART-LERECO. Je peux affirmer sans l'ombre d'un doute que l'expérience de la thèse n'aurait pu être la même sans vous. À ceux qui nous ont inspiré et prouvé que c'était possible : François, Julia, Romain, Youcef. À nous quatre qui avons commencé en même temps et qui nous sommes prouvés à nous-même que c'était possible : Francesco, Ibirénoyé, Lucile et moi-même. Et aux doctorants actuels qui tiennent bon : Ahmet, Fanny et Marie. Une petite pensée également pour les CDD qui, même s'ils n'étaient pas assez fous pour faire une thèse, n'en ont jamais eu marre de nous entendre nous plaindre. À moins que...

Je n'oublie pas nos collègues de la faculté d'économie de Rennes. Une mention toute particulière à la SDD-ESR, qu'il est bon de savoir que d'autres vivent la même chose que nous.

Je voudrais également remercier les membres de l'équipe AECP pour leur accueil lors de mon séjour à l'ETH Zürich. Je pense tout particulièrement à Niklas, mon co-bureau et co-auteur. Mais également à Ladina et Karin, les accros du spinning ainsi qu'à Yanbing, qui fût également une formidable partenaire sportive. Et comment ne pas mentionner Giulia, ma colocataire lors de ce séjour suisse, et bien plus encore.

À Youssef, mon collègue ATER et co-bureau à l'ENSAI, qui a eu la patience de m'écouter râler et qui a fait de son mieux pour m'encourager. Aussi bref que ce fût, conditions sanitaires obligent, je n'aurais pu imaginer meilleur co-bureau !

À la principale garante de ma santé mentale, à celle qui a fait d'Upside Down Rennes une deuxième maison remplie de bienveillance où chaque personne peut s'exprimer au travers de la pole dance, du cerceau aérien ou de l'exotique dance. Cette découverte sportive et artistique a bouleversé ma vie. Alors un grand merci à toi Amandine de continuer à nous transmettre ta passion. J'espère avoir un jour l'occasion de partager la scène avec toi, tous niveaux de folie acceptés !

À ceux dont j'avais croisé le chemin bien avant de commencer cette thèse et dont le soutien indéfectible est précieux pour moi. Je pense notamment à Almira, Emeline, Nathan et Thibaud. Et à ceux, dans la scène rennaise, qui sont devenus des personnes chères à mon cœur : Lucile, Enora et Corentin, ainsi que mon cher Kévin. Je chéris les moments partagés à vos côtés et espère pouvoir en partager d'autres à l'avenir. Dans le cas contraire, merci de m'avoir supporté toutes ces années, une tâche que je sais complexe !

Et à ceux pour qui de simples **mots** ne suffirait à exprimer toute l'étendue de ma reconnaissance. *À vous, à vous, à vous...*

Contents

Remerciements	vii
General Introduction	1
0.1 Agronomic principles and brief history of “Low Input” CMPs	6
0.1.1 “Low-Input” CMPs as induced technological innovations	6
0.1.2 Recent trends in input and wheat prices, and their potential impacts on “Low-Input” CMPs	8
0.2 Production function, technology and endogeneity	10
0.3 Recovering latent technology	13
0.4 Thesis outline	15
1 A multistep estimation approach for primal production functions account- ing for technology choice: a comparative evaluation	19
1.1 Introduction	19
1.2 A review of the techniques to assess the effects of production tech- nologies on farmer choices and production outcomes	21
1.3 Standard approaches to account for agricultural technology choices	23
1.3.1 Considering a technology shifting effect in the yield supply and/or input demand models	24
1.3.2 Considering technology as a regime associated to specific pro- duction and input demand models	26
1.3.3 Technology specific production function in a panel data ERS model	31
1.4 A new modelling framework to account for technology choice in agri- cultural production using the primal	32
1.4.1 Technology specific yield and input demand functions in a ERS framework	33
1.4.2 Insights for the estimation procedure	34
1.4.3 Identifying assumptions, model nonlinearity and other estima- tion methods	38
1.5 Discussion	40
1.6 Appendices	41
1.6.1 Calculation details	41

2	Estimation of production functions in low- and high-input production practices	43
2.1	Introduction	43
2.2	Background	45
2.2.1	The economic and agronomic principles of low-input production system	45
2.2.2	How to integrate technology in farmers' production function?	46
2.3	Methodological framework: from a standard ERS to an "extended" ERS model	48
2.4	Econometric implementation	49
2.4.1	A Cobb-Douglas crop production function with a damage abatement part	49
2.4.2	An original, multistep, estimation procedure	50
2.4.3	Robustness checks	53
2.5	Data	54
2.6	Insights from the technology choice and the input uses equations	57
2.7	Production function estimation results	62
2.8	Discussion and Conclusion	68
2.8.1	The input endogeneity issue	68
2.8.2	General considerations	68
2.8.3	Conclusion	68
2.9	Appendices	70
2.9.1	Results from the separability and asymmetry tests	70
2.9.2	Additional descriptive statistics	75
2.9.3	Estimation results from yield offer functions	76
2.9.4	Estimation results when considering the TFI pesticide indicator	77
2.9.5	Results when using global pesticide use variable in the production function	80
2.9.6	Results when using revenue instead of yield in the production function	81
2.9.7	Results when using a Translog specification for the productive part	82
3	Crop production models accounting for latent CMP choices	83
3.1	Introduction	83
3.2	Agronomic principles and brief history of "Low Input" CMPs	87
3.3	A Random Parameter Hidden Markov Model for modelling production choices accounting for CMPs	89
3.3.1	Crop production models accounting for CMP choices	89
3.3.2	Latent CMPs models	90
3.3.3	A model with dynamic CMP choice	92
3.4	Sketch of the estimation procedure	95

3.5	Data	98
3.6	Insights from “exploratory” analyses	102
3.6.1	Presentation of the exploratory approaches	102
3.6.2	Insights from results and comparison with the HMM approach	104
3.7	Results	110
3.7.1	Random parameters ex-post distribution	110
3.7.2	Characteristics of the three CMP categories	113
3.7.3	Simulations results	120
3.8	Discussion and Conclusion	122
3.9	Appendices	126
3.9.1	Detailed estimation procedure	126
3.9.2	Insights about k-means and AHC algorithms	132
3.9.3	Additional results for the "exploratory analyses"	133
3.9.4	Distribution of random parameters from RPHMM	140
3.9.5	Simulation results	143
3.9.6	Determinants of input uses and yield	145
	General Conclusion	147
	Bibliography	151

List of Figures

1	The successful diffusion of innovation according to Rogers, 1962	4
2	Average farmgate wheat nominal price, France, 1990-2018	8
3	Price index ratios, inputs to wheat, France, 1990-2016 (1 in 1990)	9
1.1	Intercept versus slope change in production functions accounting for technology choices	25
3.1	Annual mean fertilizers expenses and nutrients uses from 1999 to 2014	100
3.2	Annual mean yields and input expenses from 1999-2014	101
3.3	Wheat, pesticide and nitrogen annual prices from 1999-2014	101
3.4	Estimated annual mean yield of high-yielding and low-input farmers, from 1999 to 2014	105
3.5	Estimated annual mean input uses of high-yielding and low-input farmers, from 1999 to 2014	106
3.6	Estimated share of farmers who adopted a low-input CMP, from 1999 to 2014	107
3.7	Estimated annual mean yields of high-yielding, intermediate and low- input CMPs, from 1999 to 2014	108
3.8	Estimated annual mean input uses of high-yielding, intermediate and low-input CMPs, from 1999 to 2014	109
3.9	Annual mean yields for the three CMP categories obtained with RPHMM, from 1999 to 2014	113
3.10	Annual mean input uses for the three CMP categories obtained with RPHMM, from 1999 to 2014	114
3.11	Annual mean fungicide uses for the three CMP categories obtained with RPHMM, from 1999 to 2014	115
3.12	Annual mean herbicide uses for the three CMP categories obtained with RPHMM, from 1999 to 2014	116
3.13	Estimated annual share of farmers who adopted a low-input CMP, from 1999 to 2014	117
3.14	Estimated annual expected return for each CMP type (a) <i>versus</i> ob- served annual wheat prices (b)	119
3.15	Annual change in the CMP adoption share after simulating a 100% tax on chemical inputs	120
3.16	Annual change in the CMP adoption share after simulating a 10% price premium for low-input wheat producers	121

3.17 Latent class model with k-means initials – Annual mean yield in high-yielding and low-input CMPs	134
3.18 Latent class model with k-means initials – Annual mean input uses in high-yielding and low-input CMPs	135
3.19 AHC clustering – Annual mean yield in high-yielding and low-input CMPs	136
3.20 AHC clustering – Annual mean input uses in high-yielding and low-input CMPs	137
3.21 K-means clustering – Annual mean yield in high-yielding and low-input CMPs	138
3.22 K-means clustering – Annual mean input uses in high-yielding and low-input CMPs	139
3.23 Distribution of cost parameters from the technology choice models . .	140
3.24 Distribution of scale parameters from the technology choice models . .	141
3.25 Distribution of the input use and output levels parameters	141
3.26 Distribution of discount parameters from the input use and output level models	142
3.27 Annual change in the CMP adoption share after simulating a 50% tax on chemical inputs	143
3.28 Annual change in the CMP adoption share after simulating price premiums for low-input wheat producers	144

List of Tables

2.1	Descriptive statistics for low- and high-input Swiss wheat farmers . . .	55
2.2	Mean estimates of the coefficients from the Probit model for technology choice	57
2.3	Results from detailed pesticide uses models for low- and high-input production practices	59
2.4	Results from nitrogen uses model for low- and high-input production practices	60
2.5	Bootstrapped average treatment effect and the 95% confidence interval on the logarithm of input use levels	61
2.6	Results from different production function specifications accounting for production practices	63
2.7	Bootstrapped estimated yield levels (q/ha) and its 95% confidence interval for the Dummy, Separate and ERS specifications	65
2.8	Bootstrapped estimated yield levels (q/ha), or revenue (CHF/ha), and its 95% confidence interval obtained for the different robustness checks	66
2.9	Results from separability tests	70
2.10	Asymmetry robustness check - Production function estimates with different interactive inputs	71
2.11	Results from the asymmetry test	73
2.12	Symmetry robustness check - Production function estimates with symmetry of inputs	74
2.13	Repartition of Low-input / High-input farms among years	75
2.14	Results from yield offer model for low- and high-input production practices	76
2.15	Results from detailed pesticide uses models for low- and high-input production practices, with TFI pesticide indicator	77
2.16	Results from different production function specifications accounting for production practices, with TFI indicator for pesticide use	78
2.17	Bootstrapped estimated yield levels and its 95% confidence intervals obtained with our Dummy, Separate and ERS specifications, with TFI indicator for pesticide use	79
2.18	Global pesticide robustness check - Production function estimates with global pesticide indicator	80
2.19	Revenue robustness check - Production function estimates with revenue as the explained variable	81

2.20	Translog robustness check - Production function estimates with Translog specification for the productive part	82
3.1	Annual mean (and standard deviation) of arable land area and wheat acreage, from 1999 to 2014	99
3.2	Estimation standard errors, mean and standard deviation from the ex-post distribution of the CMP choice probability functions random parameters	111
3.3	Estimation standard errors, mean and standard deviation from the ex-post distribution of random parameters from the yield and input use models	112
3.4	Relative prices as non-determining factors of input and yield mean levels	145

List of Abbreviations

2SLS	Two-Stage Least Squares
AEP	Agri-Environmental Program
AHC	Ascendant Hierarchical Clustering
AIC	Akaike Information Criteria
BIC	Bayesian Information Criteria
CAP	Common Agricultural Policy
CMP	Cropping Management Practices
DiD	Difference-in-Differences
EM	Expectation-Maximisation
ERS	Endogenous Regime Switching
EU	European Union
GMM	General Method of Moments
HMM	Hidden Markov Model
HY-CMP	High-Yielding Cropping Management Practices
IPM	Integrated Pest Management
LI	Load Index
LI-CMP	Low-Input Cropping Management Practices
LIML	Limited Information Maximum Likelihood
LS	Least Squares
MAP	Maximum A Posteriori
MCMC	Markov Chain Monte-Carlo
ML	Maximum Likelihood
PCA	Principal Component Analysis
RPHMM	Random Parameter Hidden Markov Model
SAEM	Stochastic Approximate EM
TFI	Treatment Frequency Index
US	United States

General Introduction

In 1962, Rachel Carson reveals to the whole American society the negative impact of pesticides on the environment in her book entitled *Silent Spring*.¹ Since then, the adverse effect of pesticides is still a matter of concern for the scientific community. Not only pesticides are affecting the biodiversity, soil composition, water quality but they are also altering human health (Pimentel and Greiner, 1997; Pimentel et al., 1992; Wilson, 1999). Despite plentiful evidence of their noxiousness, pesticides are still abundantly used in the French agriculture, and more globally in the European agriculture. In Europe, France is amongst the biggest pesticide consumer with Spain and Italy: in 2018, France is the first of the European Union (EU) countries in terms of pesticides sales with more than 80 millions kilograms.^{2,3} But France is also the country with the largest total cultivated area in Europe with more than 27 millions of hectares of cultivated area.² When accounting for the total cultivated area, pesticide sales in France is close to the European average (respectively 2.6 kg/ha in France and 2.3 kg/ha in Europe in 2016). Yet, when looking at the evolution of pesticides sales between 2011 and 2016, they increased in France either we consider the sales in kilograms or in kilograms per hectare. This increasing trend is quite recent as from the 1990s, we were rather observing a decrease in pesticide consumption. In France, in 1999, pesticides sales were peaking at 120.5 millions of kilograms and they were almost divided by two by 2011 with 63 millions of kilograms.⁴ The increase observed during the 2010s in the French pesticide consumption contradicts the 2009 European directive – Directive 2009/128/EC⁵ – whose objective is to reduce pesticide uses in European agriculture. This EU directive requires member states to implement programs in order to achieve such goal but there is no bounding obligation of result. France developed a plan in 2008 named *Ecophyto* aiming at reducing pesticide use by 50% in 10 years (*i.e.* on the 2018 horizon). Farmers are encouraged to reduce their pesticide use but there is no coercive measures. When looking at the 2018 pesticide sales, we can see that this goal is far from being achieved. Even the 25% reduction goal introduced by the second plan, *Ecophyto II*, on the 2020 horizon is not even close

¹Re-edition: Carson (2002).

²Source: Eurostat.

³The amount of active ingredient does not permit to account for the heterogeneous impact of pesticides in terms of human health or environment. There exist specific indicators, such as the Load Index, that are accounting both for the amount of active ingredients as well as the noxiousness of the considered pesticide. Yet, contrary to the amount of active ingredients, this type of indicator is scarcely available in the data.

⁴Source: UIPP, the French union of the plant protection industry.

⁵See *Directive 2009/128/EC of the European Parliament and of the Council of 21 October 2009 establishing a framework for Community action to achieve the sustainable use of pesticides 2009*.

to be reached.

Despite strong adverse effects, pesticides are still widely used in the European agriculture. The question asked by Wilson and Tisdell (2001) in their article *Why farmers continue to use pesticides despite environmental, health and sustainability costs* is still relevant. Not long ago, use of biochemistry in agriculture was limited. But in less than one century, it seems that it became an indispensable element of modern agriculture. Hence the farmer's reluctance to change their practices especially when pesticide use guarantees high yields over time. Indeed, the combination of (i) the extensive use of pesticides and artificial fertilizers, (ii) the use of productive cultivars⁶ and (iii) the generalization of mechanization and motorization allowed tremendous productivity gain. For instance, in France, wheat yield was around 15 quintals per hectare at the end of the Second World War and reached 80 quintals per hectare at the end of the 1990s. For some years now, wheat yields seem to have reach a ceiling (Brisson et al., 2010) and are more volatile across years.⁷ The productivity gain – even if stagnating this past decade – allowed by pesticides and their user-friendliness combined with their relative low price (compared to wheat or work) plead for a large use by farmers. Farmers are thereby neglecting the environmental and societal costs of pesticide use. This difference between the social and the individual costs is one of the fundamental aspects of the externality concept. Pesticide use can be seen as a **negative externality** from agricultural production as it has detrimental effect both on human health and on the environment. As long as this social cost is not internalized by farmers, they will likely keep to have excessive pesticide use compared to the socially desirable level. A classical way to account for such difference is to introduce a **Pigouvian tax** in order to equalize the individual and social costs. The introduction of a pesticide tax is supported by some economists as Lichtenberg (2004) and Aubertot et al. (2005). France is among the few European countries with Norway, Denmark and Sweden where a pesticide taxation scheme was introduced. First, this tax was part of the general tax on polluting activities that was implemented in 2000. It was a volume tax on pesticides that was to be paid by pesticide distributors. This tax lasted until 2009 and was replaced then by a licence fee for agricultural diffuse pollution. Contrary to the previous tax, this fee is to be paid by the farmer when purchasing pesticides. Yet, the amount of such tax appeared to be quite limited, especially when comparing it to the Norway and Danish cases (Böcker and Finger, 2016). More globally, the authors conclude to a limited effectiveness of such pesticide taxes.

⁶In the European Union, contrary to other countries, genetically modified organisms are not allowed. Yet, crop breeding techniques permit to select more productive cultivars.

⁷This stagnation of wheat yields is a phenomenon that is also affecting Switzerland (see Finger, 2008) and other European countries, except Germany.

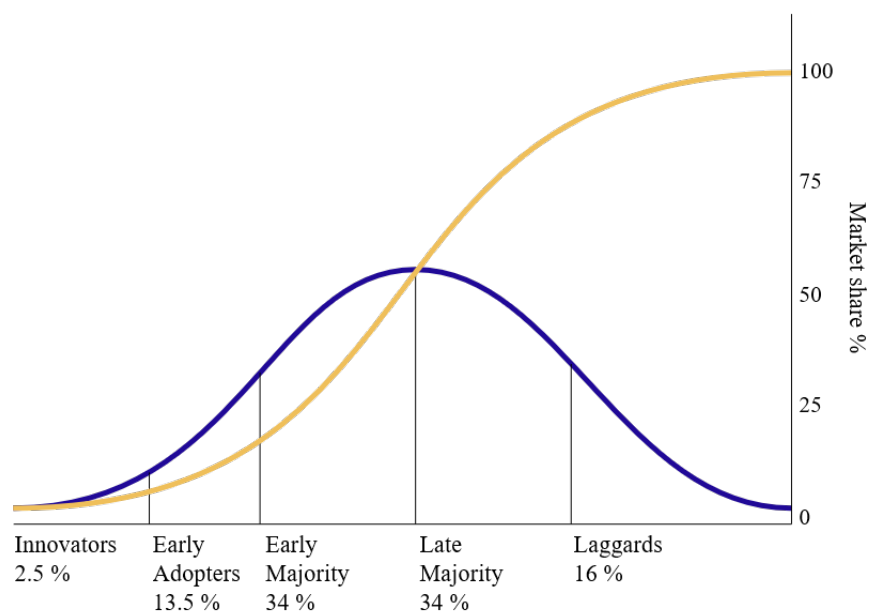
Such limited effectiveness of pesticide taxes is coherent with econometric results that are generally demonstrating that farmers' pesticide uses display very limited responsiveness to pesticide price increases (Böcker and Finger, 2017; Skevas, Lansink, and Stefanou, 2013). This provides support to the hypothesis stating current agricultural production technology "heavily depends" on pesticide uses. Yet, microeconomic analyses of farmers' chemical input uses are generally based on panel data with short time dimension and assume that farmers' production technology remains unchanged throughout the considered period. Hence, they mostly reveal that farmers' chemical input choices are inelastic given their current technology choices. When considering a conventional and low-input production functions, Féménia and Letort (2016) find that pesticide demand is more responsive to price changes. Actually, such finding highlights the fact that it is hard to reduce significantly pesticide use without a change of production technology. Indeed, pesticides and other chemical inputs are the keystone of conventional practices (Aubertot et al., 2007). As long as a farmer keeps on using those input-intensive conventional practices, his opportunity to reduce pesticide is limited. To significantly reduce pesticide use, farmers must rely on low-input techniques and integrated pest management (IPM). From that remark derives two main points. First, agronomic principles that are behind each type of practice – low-input *versus* conventional practices – being distinct, one might consider using distinct production functions (see Section 0.1). Secondly, we can state that encouraging a decrease in pesticide use is equivalent to design a program supporting the adoption of new agricultural practices that are more environment-friendly. One question is then to develop the optimal instrument to promote those pesticide saving practices. In order to do so, determinants of technology adoption should be investigated. One can also wonder how to evaluate the effectiveness of such programs either *ex-ante* or *ex-post*? In particular, will farmers react homogeneously to such programs, even if they are heterogeneous and have different farming practices? Plus, can we measure the impact of existing agri-environmental programs (AEP) – which aim at reducing the use of pesticides – by comparing the yield and input use levels observed among adopters and non-adopters?^{8,9}

⁸As for agri-environmental programs, we can take as an example the DEPHY network in France. The creation of such network is a part of the first *Ecophyto* plan and consists of a network of voluntary farmers who decided to experiment techniques to reduce their input uses. This network also mobilizes research and development resources to develop the innovations that will permit the decrease in pesticide use. The advantage of constituting a network is that each farmer can get the feedback from other farmers on experimented practices.

⁹In Switzerland, the *Extensio* program can be seen as an AEP. It is based on voluntary subscription and requires that the farmer do not use neither fungicide nor insecticides during the whole crop season. If doing so, they benefit from price premiums as well as a 400 CHF subsidy per hectare. This program is non-coercive as at every moment of the crop season they can decide to drop out and use insecticides or fungicides.

There is an abundant literature focusing on the adoption of agricultural production technologies, in economically developed or developing countries. Griliches (1957) is considered as the pioneering work initiating this literature. Yet, before the work of Griliches, we can cite the work of Tarde (1903) who describes the curve of innovation as being S-shaped. According to Griliches (1957), such shape derives from farmer heterogeneity. Indeed, when studying the diffusion of hybrid corn, he emphasizes the heterogeneous economic returns that were associated to such innovation. Such heterogeneity in returns was also highlighted more recently by Suri (2011) and Michler et al. (2018). From that perspective, early adopters are the farmers for whom the innovation is the most profitable. And, given that the cost of the innovation is expected to decrease with time, it will become profitable for more and more farmers. On the other hand, Ryan and Gross (1943) in their study highlight the social process that is behind the innovation adoption. From their study on the diffusion of hybrid seed corn, they conclude that “Commercial channels, especially salesmen, were most important as original sources of knowledge, while neighbors were most important as influences leading to acceptance”.¹⁰ This illustrates the importance of information exchanges. Another social aspect of the social process behind technology adoption is imitation – also called the learning-by-others process. The importance of imitation in the technology adoption process is emphasized by Rogers (1962) who reutilises the five stages of adoption given by Ryan and Gross (1943). The five steps are represented in Figure 1.

FIGURE 1: The successful diffusion of innovation according to Rogers, 1962



¹⁰The pesticide saving certificates implemented with *Ecophyto II* tend to turn pesticide retailers into those “original sources of knowledge”. Pesticide retailers are constrained into implementing a set of actions known for decreasing pesticide use among their buyers.

Innovators and *Early Adopters* can be seen as less risk-averse since uncertainty is an inherent characteristic of those new technologies whose returns are unknown. Once the innovation implemented by the innovators and early adopters, if the returns turn to be positive for them, others will adopt the technology by exploiting the information gathered from other farmers.

Other determinants of technology adoption were emphasised in the literature such as technical efficiency in organic farming (e.g., Kumbhakar, Tsionas, and Sipiläinen, 2009; Latruffe and Nauges, 2014), labour-constraints (e.g., Fernandez-Cornejo, Hendricks, and Mishra, 2005) or the role of risk and uncertainty (e.g., Chavas and Nauges, 2020; Marra and Pannell, 2003). Whatever the adoption determinants, the prominent fact is that adopters tend to have different characteristics from non-adopters. If such difference in characteristics is also affecting their baseline performance or their input uses, as suggested by the finding on technical efficiency of organic farming adopters, there might be **selection biases**. Then, comparing the *a posteriori* yield and input use levels of adopters *versus* non-adopters is not adequate. Such concern about potential selection biases seems to be confirmed by the finding of Finger and El Benni (2013). In Switzerland, farmers who adopted low-input practices tend to have lower yield and input use levels at baseline compared to farmers who did not adopt them. Even more, if farmers adopting low-input practices have a comparative advantage in adopting them, there might be **self-selection effects** combined to selection biases.

The main objective of the thesis is to take into account the different types of technology in production functions. Our attention will be focused on winter wheat producers in France – and in particular in *La Marne* department – and in Switzerland. Cereals in France represent more than 50% of the total cultivated area. In total, cereals in 2019 represented more than 71 millions of tonnes in France for a surface exceeding 9 millions of hectares. In terms of cereal production and surface area, *la Marne* is the first department with a respective production of 2.5 millions of tonnes and more than 0.3 millions of hectares.¹¹ Among cereals, at the national level, winter wheat represents more than 50% of the cereal cultivated area as well as the cereal production. In 2019, wheat in France represented a total of 5 millions of hectares cultivated and 39 millions of tonnes produced.^{12,13} Moreover, arable crops totalize more than 60% of pesticide expenditures due to their large acreages.¹⁴ Even if wheat crop is not the most pesticide consuming crop (the first one being potatoes), it is among the most pesticide consuming crops (after potatoes and wine) in France with a treatment

¹¹Source: SAA 2019.

¹²Source: SSP récolte 2019.

¹³Statistics on the production and surface of winter wheat are not available at the department level. This is the reason why we are not giving the figures for *La Marne*.

¹⁴Source: calculation done in Butault et al. (2011) based on Agreste data.

frequency index (TFI) of 4.¹⁵ The main contributor to pesticide uses in wheat crop are herbicides. As for Switzerland, even if the context of cereal production is quite different, it gives more coherence to consider the same crop as in France. Indeed, in Switzerland, cereals are less represented as prairies are representing more than 70% of the total cultivated area. After prairies, come cereals with 13.5% of the total cultivated area and wheat represents 50% of the cereal surface with approximately 82 000 hectares in 2019. In terms of production, Switzerland has produced around 430 000 tonnes of winter wheat in 2019.¹⁶

0.1 Agronomic principles and brief history of “Low Input” CMPs

Cropping management practices¹⁷ (CMPs) are defined by agronomists as ordered sequence of yield production decisions or decision rules aimed to produce a given crop. CMPs include soil preparation operations and type, seeding type, date and density, fertilization and pesticide applications, *etc.* We are interested in specific CMPs, the low-input CMPs (LI-CMPs) proposed by agronomists in the mid 1980s and, then, developed and promoted by agronomists and extension agents since the mid 1990s. LI-CMPs were developed by INRA starting in the mid 1980s and combined with multi-resistant wheat cultivars in the late 1990s. The “LI-CMP and hardy wheat cultivars” package were promoted by agronomists, extension agents and French wheat breeders starting in the late 1990s (Larédo and Hocdé, 2014).

0.1.1 “Low-Input” CMPs as induced technological innovations

LI-CMPs can be interpreted as technological innovations aimed to provide answers to two main issues raised by conventional high-yielding CMPs (HY-CMPs).¹⁸ First, HY-CMPs are intensive in chemical input uses, which are polluting inputs. LI-CMPs were primarily designed for reducing pesticide uses. Second, the decrease in grain prices induced by the – progressive for cereals while sudden for oilseeds – removal

¹⁵Treatment frequency index is an indicator that was built to reflect the “phytosanitaire pressure” exerted on the environment at the plot level (see Pingault et al., 2009). There is no absolute scale for TFI as it represents the number of recommended pesticide doses that were applied to the plot. In France, the “Pratiques culturales” surveys permitted to establish that, in 2001, the TFI of winter wheat as around 4, a level that is similar to sugar beet and colza. The highest TFI is associated to potatoes and is situated around 16. Interestingly enough, when comparing the results from 1994 and 2001 surveys, the TFI of arable crops in Champagne region – *La Marne* is part of this region – increased by approximately 35%.

¹⁶Source: <https://www.agrarbericht.ch/>.

¹⁷The agronomic principles of cropping management practices will be of particular interest when trying to recover unobserved technology. As extensive and intensive practices are clearly identified in Swiss data but not in France. Therefore, we will focus on France during this presentation as technology are latent in French data. Yet, the agronomic principles that are presented are applicable to Swiss extensive practices.

¹⁸In this manuscript, we use both the high-yielding and high-input denominations to refer to these conventional practices. Occasionally, we also refer CMPs as production practices.

of the CAP price support in 1992 called into question the profitability of grain production in the EU from the late 1990s to the mid 2000s. Due to the low grain prices during this period, HY-CMPs appeared to be much less profitable than they were in the early 2000s.

The price support implemented by the CAP until the so-called McSharry reform in 1992 led most agricultural scientists to develop HY-CMPs to be adopted by European grain producers. Indeed, due to the relative scarcity of arable land in Western Europe, adopting HY-CMPs appeared to be the most profitable technological option for farmers to benefit from high grain prices (Mahé and Rainelli, 1987; Meynard, 1991). HY-CMPs aim to increase grain potential yield by increasing seeding densities, choosing early seeding dates, relying on productive seed varieties and applying large amounts of, especially nitrogen, fertilizers. Importantly, these HY techniques tend to increase pest and weed pressures and, consequently, call for efficient crop protection. Early seeding dates tend to expose crops to pest outbreaks. Nitrogen fertilizer uses tend to trigger competition by weeds. High seed densities, productive – but susceptible to diseases – cultivars and high loads of nitrogen fertilizer tend to increase wheat susceptibility to diseases. Yet, availability of efficient, as well as relatively cheap, chemical pesticides enable farmers to control the pest and weed pressures triggered by HY techniques.

LI-CMPs were conceived by agricultural scientists as an agronomic response to the polluting emissions induced by the use of chemical inputs, of chemical pesticides in particular, and to the decrease in grain prices due to the CAP reform implemented in 1992 – the so-called McSharry reform. HY-CMPs are conceived to achieve high target yield levels but rely on high levels of chemical input uses, precisely because the techniques implemented for achieving high target yield levels tend to trigger the need of high fertilization and crop protection levels. The basic principle of the conception of LI-CMPs is to slightly lower target yield levels in order to significantly lower chemical input uses, pesticides in particular. Lowering target yield levels directly reduces crop nutrition needs and, thereby, nitrogen fertilization uses. LI-CMPs reduce crop protection needs by avoiding cropping techniques that increase pest and weed pressures. Therefore, they allow reducing pesticide uses.

The HY-CMPs and LI-CMPs considered by agronomists vary across time and production areas, depending on economic and agro-climatic conditions (Bouchard et al., 2008; Loyce and Meynard, 1997; Loyce et al., 2008, 2012; Rolland et al., 2003). On average, the yield levels obtained with LI-CMPs are 10% lower than those obtained with HY-CMPs. Nitrogen fertilizer loads decrease by 10% from the HY-CMPs to the LI-CMPs while the use of (mostly) fungicides and insecticides is reduced by around 30%. Finally, due to the lower sowing densities in LI-CMPs seed uses decrease by

around 50% when using these CMPs. Also, hardy wheat cultivars are complementary to the agronomic principles underlying the design of LI-CMPs (Larédo and Hocdé, 2014; Loyce et al., 2008). These cultivars are resistant to multiple diseases but slightly less productive than the ones typically used in HY-CMPs. Finally, LI-CMPs are labor and fuel saving thanks to their lower expected pesticide application numbers.

0.1.2 Recent trends in input and wheat prices, and their potential impacts on “Low-Input” CMPs

Due to the CAP price support, the wheat price paid to French farmers was relatively high until the McSharry reform 1992 (see Figure 2). Then it progressively declined and attained the world price level in the late 90’s. World wheat price remained low until 2006. It dramatically increased in 2007, following the US biofuel policy and increasing fuel and feed demands. Wheat price has remained relatively high on average since then, though it is volatile due to random shocks on cereal supply and demand.

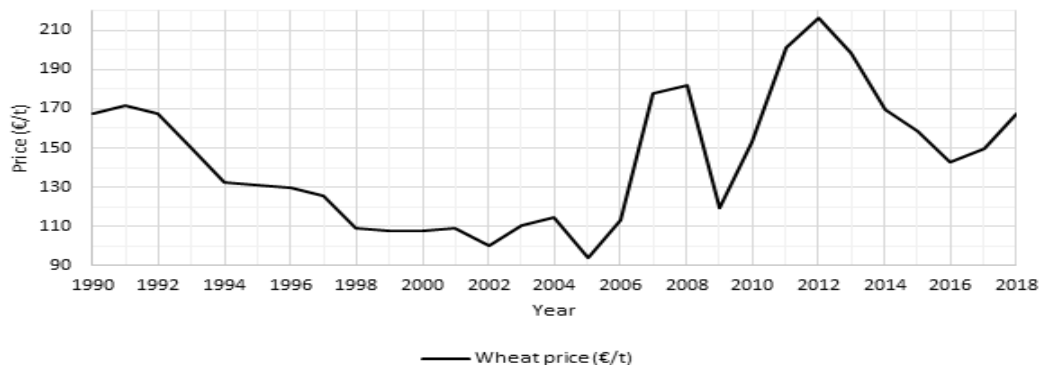


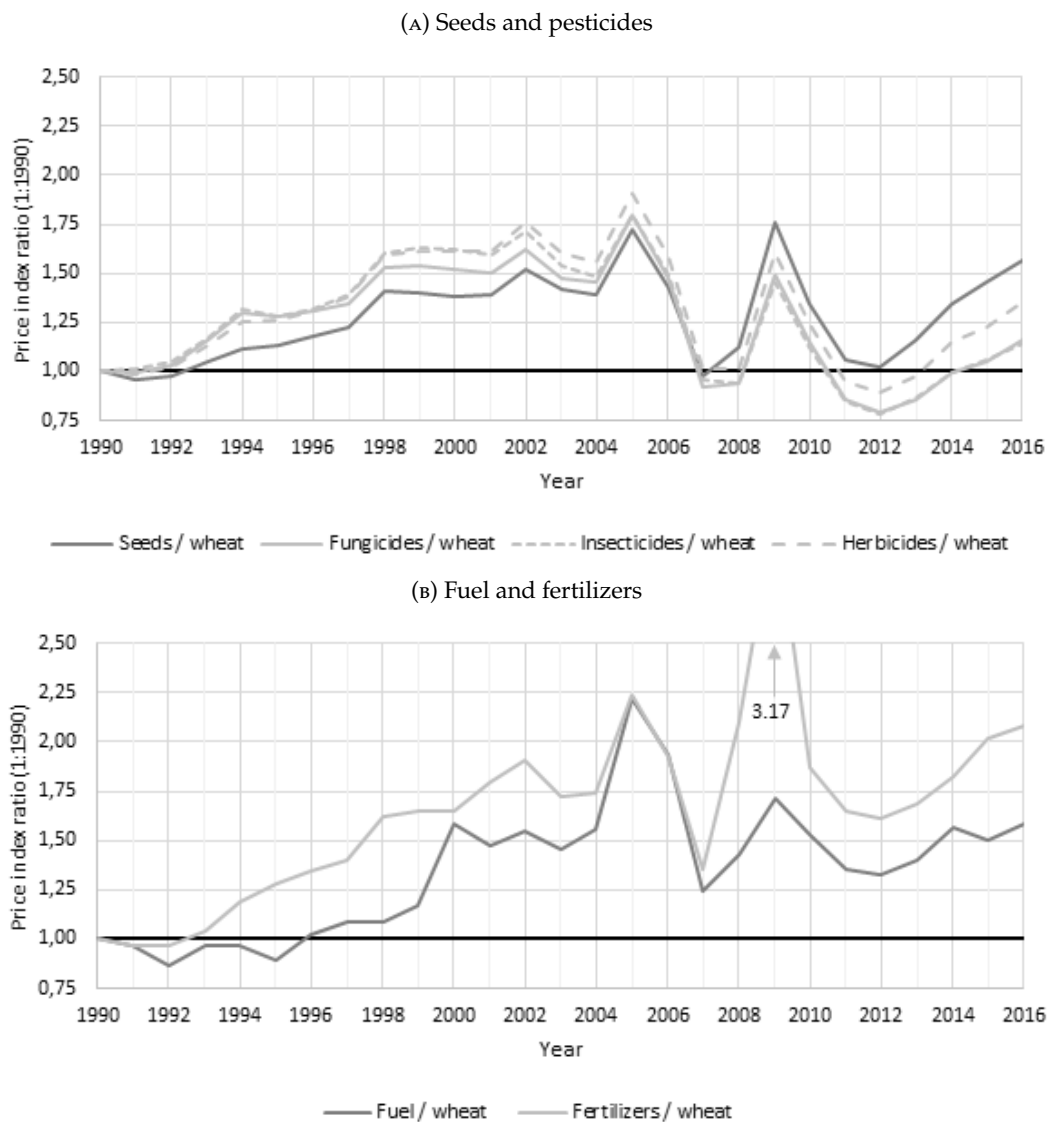
FIGURE 2: Average farmgate wheat nominal price, France, 1990-2018

The early 1990s are considered as the peak of the chemical intensity of the French arable production, with price ratios favoring the use of chemical inputs and, as a result, of HY-CMPs. LI-CMPs were developed starting in the mid 1980s. They started to be tested on farm fields and promoted in the 1990s, as responses to the low wheat prices observed from the late 90s to 2006. The yield reductions entailed in adoption of LI-CMPs have much less detrimental effects on profits when crop prices are low. But, the high wheat price levels observed since 2007 suggest that HY-CMPs might be as profitable today as they were in the early 1990s. Yet, input prices increased since the early 1990s, at difference rates though.

Figure 3 displays the evolution of the ratio of the input price indices to that of wheat from 1990 to 2016, using 1990 as the baseline. Basically, an input is cheaper (more

expensive) for wheat production in year t than it was in 1990 when ratio of the price index of this input to that of wheat is below (above) 1. Panel (A) in Figure 3 shows that pesticide prices were about 50% more expensive relative to wheat in the early 2000s than they were in the early 1990s. Since 2007, the ratios of the prices of pesticides to that of wheat have been only slightly above the ones observed in the early 1990s. Indeed, the nominal prices of fungicides and insecticides remained steady while that of herbicides increased only by 15% from 1990 to 2016.

FIGURE 3: Price index ratios, inputs to wheat, France, 1990-2016 (1 in 1990)



Fuel and fertilizers increased significantly during the 1990s, as shown by panel (B) in Figure 3. Fertilizers were 75% more expensive relative to wheat during the 2000s and 2010s than they were during the early 1990s. The corresponding ratio is around 50% for fuels. Differences in input uses between LI-CMPs and HY-CMPs are less important for fuel and fertilizers than they are for fungicides and insecticides. Yet,

fuel and fertilizers price levels were substantially higher from 2007 to 2016 than in the early 1990s.

This factual analysis suggests that economic conditions tended to favor adoption of LI-CMPs from the late 1990s to 2006, mostly due to the low grain prices observed during this period. The prices of fungicides and insecticides, the use of which is aimed to be reduced by adopting LI-CMPs, remained stable from 1990 to 2016. The high grain price levels observed since 2007 have tended to favor conventional HY-CMPs, although these effects of high grain prices on the profitability of HY-CMPs are partially offset by the high levels of fuel and fertilizers prices.

0.2 Production function, technology and endogeneity

The presence of technological heterogeneity among farmers is a fact not to be discussed. What need to be discuss however is how to account for such heterogeneity in production functions. A naive approach would be to introduce a dummy variable when technology is observed. By doing so, one only allows for a shift in yield and input use due to technology choice and induce homogeneous reaction to public policies or changes in input prices. The differences between low-input and high-yielding CMPs described in the previous section suggest that different CMPs need to be considered as different crop production technology. For instance, production functions describing high-yielding CMPs are expected to be more responsive to fungicide uses than those describing low-input CMPs. Low-input CMPs for wheat rely on early and light sowing and on seed varieties that tend to decrease disease risks. Similarly, production functions describing high-yielding CMPs are expected to be more responsive to nitrogen uses than those describing low-input CMPs. High-yielding CMPs are, by definition, designed for achieving high yield levels that, in turn, require high fertilization rates. This variety of responses to input uses cannot be accounted for by simply allowing for a “shift” in yield level. Accordingly, different CMPs need to be described by different, CMP specific, production functions. One can argue that, instead of separate production functions, introducing interactive terms with the adopted technology might be enough as done by Serra, Zilberman, and Gil (2008). Yet, contrary to separate functions, the use of a unique model implies that the distribution of error terms is the same across technology. Such assumption is quite strong in a CMP context as it means that yield variability is similar whatever the implemented technology. The use of separate functions seems more adequate, especially given the fact that the findings from the production frontier literature show significant differences in terms of input productivity and overall efficiency depending on the farmer technology.

By using distinct production functions, one allows for heterogenous responses to input price changes or public policy more generally. The literature considering separate conventional and organic functions is particularly extensive. For instance, Gardebroek, Chavez, and Lansink (2010) estimate separately the production function as developed by Just and Pope (1978) for conventional and organic farmers. A similar approach was adopted by Lansink, Pietola, and Bäckman (2002) in a stochastic frontier analysis framework and by Gardebroek (2006). From those separate functions, one might derive yield and productivity differences between organic and conventional farming. Yet, by doing so, one implicitly assumes that organic farmers would have the same average productivity as conventional farmers if switching to conventional practices, other heterogeneity factors being controlled for (*e.g.*, farm size or other farm(er) characteristics that might affect yield). Pietola and Lansink (2001) study on the impact of the Finnish policy to encourage the adoption of organic farming practices shows the opposite. Farmers who switched to organic farming tend to already have lower performances, hence lower yield level. If all the factors that are affecting both production choices and technology adoption are observed by the econometrician, then this selection issue can easily be controlled for. Yet, some factors that are impacting both choices might be unobserved, thus leading to **selection biases**. Then, one cannot genuinely compare the performances of organic *versus* conventional farmers to derive the differences in terms of yield and input use level between both production practices.

This discussion about whether these factors are totally or partially observed by the econometrician is crucial. Indeed, it permits to better understand the controversy that arise on the ability of organic agriculture to feed the world (see the answer of Connor (2008) to the article of Badgley et al. (2007)). On that matter, Seufert, Ramankutty, and Foley (2012) argue that the yield differences between organic and conventional agriculture are “highly contextual”. Behind that argument is the idea that unobserved factors underlie lower productivity of organic farms. It advocates in favor of selection biases between conventional and organic agriculture that might be generalizable to low-input practices. In line with findings from Pietola and Lansink (2001) on organic farming, Finger and El Benni (2013) show that, after being subsidized by the Swiss authorities, low-input practices were adopted by farmers that tend to have lower yield and input use levels at baseline. This argues in favor of the presence of selection issues between low- and high-input production practices. One can also think about a **self-selection effect**. Self-selection effect is a form of selection issue arising if farmers adopting low-input practices tend to have a comparative advantage when adopting them, *i.e.* they are benefiting from this technology. This self-selection effect is in line with the innovation adoption theory of Griliches (1957). It remains that, whether they impact the economic return of each technology, the unobserved factors that impact both production choices and technology adoption need to be accounted for. Otherwise, when evaluating the impact of a technology

on production outcomes – or more generally the effect of an AEP, estimates would suffer from selection biases.

The approach proposed by Heckman (1979) to account for selection biases became standard. The method he developed to correct this endogeneity issue is to introduce a control function in the model. This control function – the inverse Mills ratio – accounts for the expectation difference and comes from a second equation called the selection equation. As initially formulated, this method was implemented for cases where only the “treated” group is observed. Lee (1982) generalizes Heckman’s approach to the case where both the treated and untreated (or control) groups are observed. This method is known as the endogenous switching regression model and also relies on the use of the inverse Mills ratio. Heckman’s correction for selection bias is widely used in the literature and was also used to control for technology endogeneity in agriculture. In particular, Greene (2010) extends Heckman’s correction to the framework of stochastic frontier models. The developed methodology was then applied by Bravo-Ureta, Greene, and Solís (2012) to evaluate the impact of a program implemented in Honduras on technical efficiency. Endogenous regime switching (ERS) models are also applied to production choice models. For instance, Alene and Manyong (2007) use an endogenous regime switching model to evaluate the impact of farmer education on the farm productivity considering both a traditional and improved technology. Asfaw et al. (2012) use this approach to evaluate the impact on technology adoption on household whereas Abdulai and Huffman (2014) estimate the impact on both yield level and economic returns.

Yet, most of time when considering production functions, other endogeneity issues tend to arise in addition to the one caused by self-selection effects and/or selection biases. In particular, production function estimation generally suffers from input endogeneity issues. While acknowledging the pioneering work of Haavelmo (1943), Marschak and Andrews (1944) were the first to formalize the input endogeneity issue in production function estimation problems. If one, reasonably, assume that farmers are economically rational then they seek to maximize their expected profit yields by optimizing their input uses. If the simultaneity of input choice and expected profit is neglected, the estimated production function is biased. The solution proposed by Marschak and Andrews (1944) to solve such input endogeneity issue is to estimate simultaneously the production function and input demand equations. An alternative approach consists in estimating the production function while instrumenting input uses – generally by input and crop price levels (*e.g.*, Mundlak, 1996). Similar approaches can be implemented in an endogenous regime switching framework. For instance, an extended approach of Heckman’s selection model is developed by Schwiebert (2015) to include endogenous covariates whereas Kumbhakar, Tsionas, and Sipiläinen (2009) consider a simultaneous equation framework to solve both covariates endogeneity and selection issues.

0.3 Recovering latent technology

Unfortunately, no data exists on the adoption of LI-CMPs by French farmers. Moreover, farm accountancy data, even with cost accounting, does not contain any indicator enabling us to identify farmers using LI-CMPs. For instance, seed cultivars are not reported. Similarly, if purchased seed expenditures vary with sowing densities, these expenditures may also vary with seed prices and the share of seeds produced farmers themselves. This information lacking implies that we can only consider inferring farmers' CMPs from their yield and chemical input use levels, that is to say indirectly. Different approaches can be considered to reveal latent structure in a data set.

Among those approaches is clustering analysis. Clustering is an unsupervised learning method whose aims at finding structure in a data set or, to say it differently it is "the art of finding groups in data" (Kaufman and Rousseeuw, 2009). Clustering analysis results in defining several groups among the data. To be meaningful, the obtained partition must separate observations that are very dissimilar while gathering similar observations. Clustering algorithms are thus trying to minimize the within group distance while maximizing the between-group distance. There exists different metrics to measure distance as well as different partitioning algorithms. Hence the need to generally compare the results obtained with different algorithms and metric to get robust results.

Clustering analysis is used in a lot of research areas and among them is agriculture. It has been used mainly to build farm classifications. For instance, it has been used by Gebauer (1987) to build a typology of farm households in Germany. Such type of typology was also used by Perrot (1990) in France and Maseda, Diaz, and Alvarez (2004) in Spain. Those typologies were foremost used to describe the state of the agricultural sector at a specific time. Clustering analysis was also used to identify different cultural systems by Bellon, Lescourret, and Calmet (2001) and Renaud-Gentié, Burgos, and Benoît (2014). Clustering is used here to summarize the detailed agronomic information available for each farm to build homogeneous groups.

Another method to recover latent variable is to use a latent class model (see Bartholomew, Knott, and Moustaki, 2011). Whereas clustering analysis assigns each observation to a class, latent class models rather estimate a probability to belong to each class. Given the number of classes we expect to have, we use a mixture model to estimate (i) the *a posteriori* probability to belong to each group and (ii) the parameters of the dependent variable distribution. Generally, the estimation of the mixture likelihood is performed with the Expectation-Maximization algorithm. In fact, latent class models can be seen as a probabilistic clustering where the distance metric is substituted by the likelihood that evaluates the model consistency. Plus, contrary to clustering methods that require a second step to estimate the separated production

functions, latent class models have the merit to compute both the production functions and the CMP affiliation in one step. In that case, this one step procedure has advantages over the two step procedure because it avoids (i) biasing the estimates of the second stage equation if some errors are affecting the first stage and (ii) losing information. The latest point is emphasized by Orea and Kumbhakar (2004): the information that is contained within a class is not used to estimate the technology of other classes. This is the very reason why, in their work, Orea and Kumbhakar (2004) prefer to use a latent class framework to estimate different production frontiers.

The work of Orea and Kumbhakar (2004) was conducted on Spanish banks and proved that latent class models allow to control adequately for unobserved heterogeneity. Latent class frontier models were also used in the agricultural literature by Alvarez and Corral (2010), Martinez Cillero et al. (2018) and Renner, Sauer, and El Benni (2021) for instance. Alvarez and Corral (2010) distinguish intensive from extensive dairy farm. As for Martinez Cillero et al. (2018), they estimated three different categories among Irish beef farms. Renner, Sauer, and El Benni (2021) consider stochastic frontier with latent class so to evaluate the impact of technology choice and change on productivity and efficiency of Swiss dairy farms. If production frontiers are found to be significantly different across technology, no doubt that this technological heterogeneity also needs to be accounted for when considering production functions.

The pitfall of the clustering and latent class approaches as described before is that it requires restrictive assumptions on the technology choice across time. Indeed, when having panel data at your disposal, the previous methods entail supposing either (i) that farmers have the same technology for each year (e.g., Alvarez and Corral, 2010; Martinez Cillero et al., 2018) or (ii) that farmers can change technology each year (e.g., Dakpo et al., 2021; Orea, Perez, and Roibas, 2015). Allowing technology change each year might be problematic. Even if CMP choices share more similarities with crop variety choices – usually considered as short run choices and modelled as such (e.g., Michler et al., 2018; Suri, 2006) – than with irrigation technologies – modelled as investment choices (e.g., Genius et al., 2014) – existence of transition costs makes frequent changes not appealing. Yet, technology change needs to be allowed, especially when considering short run decisions. Hence, the model needs to allow the possibility of change while making it costly so that individuals cannot switch every year from technology to technology. Such modelling effort was made in Dakpo et al. (2021) by considering explicitly a time trend among the variables that are used to separate the data in different groups. However, the dynamics of technology adoption is not explicitly modelled. If one is interested in the determinants of technological change to evaluate the impact of a public policy, this approach is not appropriate. A solution is to model the dynamics of technological adoption as a Markov process. Markov models are particularly adapted to model situation where

there can be regime switching every year. They were introduced in econometrics by Goldfeld and Quandt (1973) and then largely diffused with the Markov-switching regression model used by Hamilton (1989). The Markov-switching models defined by these authors are based on a latent variable of which we want to model the dynamics. They were used in economics for example to explore the dramatic breaks that can occur in economic time series (see Chauvet and Hamilton, 2006; Hamilton, 1989). It is of special interest in financial economics as abrupt changes are common in financial data (see, *e.g.*, Bonomo and Garcia, 1996; Cecchetti, Lam, and Mark, 1988)). Markov models are also used in the agricultural literature. For instance, Benoît, Le Ber, and Mari (2001) use hidden Markov model (HMM) to study the dynamics of crop rotation. Markov chain were also implemented by Miller et al. (2017) to investigate the adoption path of precision agriculture technologies in farms. The advantage of Markov models is their flexibility: transition matrices can either be (i) time invariant or change across time, (ii) be arbitrary fixed or can be defined by a logistic model.

0.4 Thesis outline

As stated previously, the main objective of the thesis is to estimate and to compare technology specific production functions. We attempted, in this general introduction, to stress the need for separate functions to account for heterogeneous technology. Yet, because technology adoption might be subject to selection biases, estimating technology specific production functions is not enough. A standard approach to deal with this technology selection issue in agricultural production is the endogenous regime switching (ERS) framework. In Chapter 1, we present the standard ERS framework as well as an extension to the case with endogenous covariates. We also present an associated estimation procedure to this ERS extension. Such extension and its estimation procedure are of particular interest in Chapter 2. As a matter of fact, in this chapter we intend to study how yields respond – and to what extent the response differs based on the adopted technology – to changes in variable input uses, especially pesticide uses. The perspective adopted in Chapter 3 is slightly different. We still want to compare CMP specific production functions, but in an unobserved technology choice context. Thus, we need to develop a framework that permits to uncover the CMP choice.

The second chapter of this thesis is mainly a theoretical chapter. We start by presenting the standard approaches to account for technology in agricultural production functions. In particular, we present the endogenous regime switching framework that permits to account for technology selection issues. Chapter 1 contribution lies in our considering an extension of the ERS framework for the endogenous covariates case. We consider an easily tractable control function approach with two sets of

control functions. The first one controls for endogenous sample selection issue associated to technology choice whereas the second controls for input use endogeneity. In particular, we show in this chapter that the expression of the so-called inverse Mills ratio used in the widely used Heckman's two step approach needs to be adapted in ERS models when regressors are endogenous.

The extended ERS model we present in Chapter 1 as well as its estimation approach directly derives from the research question we try to answer in Chapter 2. In this chapter, we intend to estimate and compare the production functions of low-input and high-input farmers. In particular, we want to study how yields respond to input uses with a focus on pesticides. This objective encourages us to consider (i) a primal production function with (ii) a damage abatement function so we can consider the protective role of pesticides. The primal function comes with the well-known input endogeneity estimation issue whereas the damage abatement function entails non-linearity estimation issues. Both issues are tackled with the approach proposed in Chapter 1, *i.e.* a multi-step estimation approach relying on control functions. Thus, Chapter 2 is an empirical application of the extended ERS model and estimation technique we developed in Chapter 1.

The empirical analysis presented in Chapter 2 uses rich, high resolution panel data on Swiss wheat production (617 observations, from 2009 to 2015), containing detailed information on output and input uses (*i.e.* on pesticide uses, mechanical weed control, fertilizer uses, work and machinery), obtained from field journals. We make use of the fact that parallel to conventional wheat production, a low-input ("Extensio") wheat production system exists in Switzerland. Thus, we can compare the conventional and low-input production functions, in particular how yield responds to input uses in both functions. Unfortunately, our empirical analysis suffers from weak instruments in the input use equations. It means that the estimation results of the production functions have to be interpreted cautiously as we do not properly account for the input use endogeneity. Despite that pitfall, this article seems to confirm the presence of selection biases affecting both the input demand and yield level models. In particular, unobserved characteristics of high-input farmers seem to boost up their pesticide use levels whereas unobserved characteristics of low-input farmers seem to allow them to reach greater yield levels. Even if estimates from the production functions need to be considered cautiously, this chapter argues in favor of controlling for selection biases when estimating input use and yield level functions and when evaluating the treatment effect of low-input adoption on both input use and yield levels.

Finally, Chapter 3 aims at proposing statistical and micro-econometric approaches for uncovering CMPs used by farmers when these are not available in the data set. Given that intensity in the use of chemical inputs of a CMP is directly related to the yield level targeted by this CMP, our methods aim to identify CMPs used by farmers based on their yield and chemical input use levels, information that are generally available in most cost accounting data set. We first considered more exploratory approaches such as clustering and latent class models to (i) identify the CMP classes and (ii) estimate CMP specific production functions. Yet, when considering panel data, those approaches suffer from their static perspective. We need either to suppose that technology is stable on the whole period, an assumption that is problematic when the size of the panel is increasing. Or, we need to suppose that the technology choice is independent from the technology observed at the previous period. This second assumption is as unsatisfactory as the first one. We need to consider the adoption process as a dynamic one. We will then assume that CMP choices can be modelled as a Markovian processes and use a hidden Markov chain model to estimate CMP specific production functions. Another aspect we integrate in our modelling framework is farm unobserved heterogeneity, which can affect both crop production and farmer choices. We account for the unobserved farm heterogeneity by considering a random parameter model. Finally, our model combines different elements that are generally considered separately: latent technologies, unobserved heterogeneity and the dynamics of CMP choices. Our model can thus be considered as an endogenous Markov switching model. Plus, by disclosing the effect of CMP returns, our model allows to investigate the effects of economic incentives on CMP choices, which is critical from a public policy perspective.

We illustrate our approaches by investigating French farmers' CMPs for winter wheat based on a panel dataset from 1998 to 2014 covering *la Marne*, a (highly productive) arable crop production area located in eastern France. We manage to distinguish three different CMPs: a high-yielding, an intermediate and a low-input CMPs. Even if low-input practices are associated to lower profit levels, they are still attracting an irreducible share of farmers. This shows that technology return is not sufficient to explain technology adoption and confirms the need to account for farmer preferences when studying adoption decision. In addition, our results show that, to encourage farmers with more intensive practices, price instruments on pesticide inputs are quite ineffective when wheat prices are high. This is due to the fact that profits depend more from the selling price of wheat than the input expenses that are small compared to the yield value. This confirms the previous findings on the inelastic demand of pesticides when wheat prices are high. A public policy implication is that, when crop prices are high, economic incentives targeting input prices might not be the best instrument to encourage a change in practices.

Overall, the results of this PhD thesis argue for the need to consider CMP in economists' production functions. Farmers using different CMPs tend to have different input uses and yield levels. Moreover, this CMP choice induces selection biases on both input use and yield levels. The unobserved factors implying selection biases might also impact the farmers' response to public policy. In particular, the response to public policies aiming at reducing pesticide uses might be heterogeneous across CMPs. This potential response heterogeneity should be considered by public authorities in order to adequately evaluate the impact of a specific policy, or more generally agri-environmental programs.

An other valuable input from this PhD thesis comes from our investigation into the CMP choice mechanisms. In particular, we show that the standard taxation instrument is not a sufficient incentive for low-input adoption. Differences in pesticide uses between low- and high-input CMPs do not compensate for the yield loss, especially in a context of high crop prices. On the other hand, price premiums for low-input wheat might help into encouraging farmers to adopt this more agri-environmental practice. This is the strategy that have been adopted by Switzerland, where low-input practices are much more established among wheat producers. Low-input producers benefit from both a price premium and a direct payment (400 CHF/ha). This might seem a rather expensive policy in comparison to the pesticide uses saving. The viability of low-input practices in wheat production as a high-input *versus* organic in-between might be questioned. In particular, a cost-benefit analysis of both production practices might be considered to answer that question.

Chapter 1

A multistep estimation approach for primal production functions accounting for technology choice: a comparative evaluation¹

1.1 Introduction

Historically, agricultural economics was interested in quantifying “the contribution of inputs to output variations” and evaluate the production elasticities (Mundlak, 2001). Yet, as pointed out by Marschak and Andrews (1944), direct estimation of the production function suffers from biased estimates due to inputs endogeneity. Such endogeneity issue comes from what analysts assume that farmer’s inputs choice is rational while most drivers of this choice remains unobserved to them. In that context, dual approach gain popularity in the early 1970s as a way to get around input endogeneity issue (e.g., Griliches and Mairesse, 1995). Duality relies on the assumption that farmer’s input demand answers profit maximization objective. This approach was first developed by Klein (1953) and applied by Wolfson (1958). Chambers et al. (1988) even wrote an entire book on “dual” approaches applied to production analysis. Yet, Mundlak (2001) deplores that, because it relies on a direct behavioral function for input demand, dual approach rarely questions the relationship between inputs and outputs.

As pointed out by Akerberg, Caves, and Frazer (2015), applied economists have devoted much attention to this endogeneity problem and approaches to solve it. Dual approach is only one of them. The first approach was proposed by Marschak and Andrews (1944), while acknowledging the pioneering work of Haavelmo (1943) on this topic. It consists in estimating simultaneously the production function and the input demand equation(s). An alternative to this “full information” approach is a “limited information” approach which consists of estimating the production function while instrumenting input uses by price levels (e.g., Mundlak, 1996). Another approach,

¹This chapter is written in collaboration with Alain CARPENTIER.

relying on more limited behavioral assumptions than the dual one, is proposed by Olley and Pakes (1996) or Levinsohn and Petrin (2003). Input endogeneity issue is tackled by introducing control functions that are based on the timing of input uses (e.g., Akerberg et al., 2007; Akerberg, Caves, and Frazer, 2015; Wooldridge, 2009).

In this chapter, we consider the input endogeneity issue within an endogenous regime switching (ERS) framework. Our considering of this framework follows the pioneering works of Pitt (1983) and Fuglie and Bosch (1995) so we can investigate the effect of cropping management practices (CMPs) on the wheat production and its associated input uses. In particular, we consider an ERS framework with non-linear production functions. Non-linearity entails estimation issues. It makes both the simultaneous estimation “full information” and the instrumental variable “limited information” approaches hardly tractable. We use a control function approach (see Wooldridge, 2015) relying on two sets of control functions. The first one controls for the endogenous sample selection issues implied by ERS models whereas the second one controls for input use endogeneity. Our estimation approach can be considered as an extension of the widely used two-step approach that was initially proposed by Heckman (1976, 1979) to account for endogenous sample selection and later adapted by Lee (1978) to the case of Gaussian ERS models. In particular, we show that the expression of the so-called inverse Mills ratio used in Heckman’s two step approach for estimating regression models under endogenous sample selection needs to be adapted when regressors are endogenous.

Although it relies on admittedly restrictive assumptions, the approach we propose here is easier to implement in practice than its competing alternatives – e.g., the generalized method of moments or the two-stage least squares (2SLS) approach proposed by Wooldridge (2010) – especially those based on estimating equations defined by orthogonal conditions. This point is further discussed in the chapter. After defining the notations we used in the rest of the chapter, we proceed in two steps. First, we present a brief literature review on the methods that were used to deal with the effects of technology choices on producers’ input choices and production outcome. In particular, we present and discuss the main issues that arise when considering endogenous technology choices, *i.e.* self-selection effects and selection biases, and the most common approaches to deal with them. Second, we present our modelling framework and the estimation approach for assessing the effects of CMPs on crop yield functions.

Notations

Let $i = 1, \dots, N$ and $t = 1, \dots, T$ represent respectively the farmer and year indexes from a panel dataset. Let y_{it} , respectively $x_{k,it}$, be the observed wheat yield level, respectively the use of chemical input k with $k \in K$, of farmer i at time t . Let $\mathbf{x}_{it} = (x_{k,it} : k \in K)$ be the K -dimension vector of chemical inputs use. Finally, let

r_{it} denotes the CMP choice of farmer i at time t where $r_{it} = 1$ (respectively $r_{it} = 0$) indicates that s/he chooses the low-input CMP (respectively the high-input CMP).

Relevant wheat and input prices are gathered in $K + 1$ -dimension vector \mathbf{p}_{it} .² Vector \mathbf{c}_{it} , of dimension C , collects farmer and farm characteristics that directly impact wheat production choices whatever the CMP choice whereas Q -dimension vector \mathbf{q}_{it} consists in farmer and farm characteristics that impact farmers' CMP choices without (directly) impacting their (CMP specific) wheat production choices. Vector $\mathbf{w}_{it} = (\mathbf{p}'_{it}, \mathbf{c}'_{it})^T$ stacks up the observed drivers of chemical input uses and, thereby, of yield levels, that is to say netput prices \mathbf{p}_{it} and farm(er) characteristics \mathbf{c}_{it} . Vector $\mathbf{z}_{it} = (\mathbf{w}'_{it}, \mathbf{q}'_{it})^T$ appends the observed drivers of the CMP choices, \mathbf{q}_{it} , to those of chemical input uses and yield levels. This vector constitutes the information set of the models considered below. It contains the variables that are assumed to be exogenous in these models.

1.2 A review of the techniques to assess the effects of production technologies on farmer choices and production outcomes

The impact of technologies on various agricultural production choices (*e.g.*, variable input uses) or outcomes (*e.g.*, yield levels, profit levels, production levels or production frontier) is studied in numerous studies. For instance, Fernandez-Cornejo (1996, 1998) considers the effects of integrated pest management (IPM) on profit and pesticide use levels in fruit production and viticulture in the United States. Fernandez-Cornejo and Wechsler (2012) focus on the effect of Bt (*i.e.* genetically modified organism) *vs.* high-input corn on profit and input use levels in the US. Livingston, Fernandez-Cornejo, and Frisvold (2016) consider the effect of glyphosate *versus* herbicide mixes based weed control on yields and profit levels in US soybeans and corn production. Khanna (2001) studies the effects of soil testing with or without variable rate fertilization techniques *vs.* standard fertilization practices on the productivity of nitrogen applications on corn in the United States (US). Teklewold et al. (2013) analyze the impacts of various combinations of seed choices, cropping system and tillage practices on profit and input use levels in maize production in Ethiopia. Suri (2011) considers the effect of hybrid *vs.* conventional seeds on maize yield levels in Uganda while Michler and Shively (2015) address the same topic for chickpea production in Ethiopia. Kumbhakar, Tsionas, and Sipiläinen (2009) compare the stochastic production frontiers of conventional *versus* organic dairy farms while accounting for endogeneity of the adoption of organic production practices.

²By default, unless specified otherwise, we consider column vectors.

All these studies acknowledge that (i) technology choices and (ii) production choices and outcomes are impacted by many common (or at least correlated) factors. Among these factors, some are observed by the analyst such as market prices, weather conditions or general farm(er) characteristics. As such, their effect can be controlled by explicitly incorporating them in the modelling framework. On the other hand, some factors as soil quality or farmers skills are unobserved and induce endogeneity issues that are to be accounted for. Yet, most studies assume that such endogeneity issues cannot be correctly addressed with observed variables. From that perspective, the approach adopted by Livingston et al. (2015) is a notable exception. They use a propensity score matching approach to estimate the effects of weed control practices on corn production outcomes. Yet, whether considering matching or inverse propensity score weighting techniques, estimates of the causal effects of technology choices are consistent only if these choices and the considered outcomes are independent conditionally on the control variables (*e.g.*, Angrist and Pischke, 2008; Imbens and Wooldridge, 2009; Wooldridge, 2010). Such conditional independence assumption might not hold when considering production technologies and their effects on production choices (*e.g.*, input use levels) at the farm level. Indeed, there exists factors, observed by the farmers but unobserved by the analyst, impacting farmers' technology choices as well as their input uses decision. Among such factors are soil quality, farmers' technical skills or pest outbreaks, that impact both technology and production choices as well as outcomes (*e.g.*, yields or profits) thus creating selection biases. On the other hand, technology choices are made on purpose by farmers so that it impacts their input use and yield levels through partly unobserved mechanisms, thereby inducing self-selection effects.

From that perspective, quasi-experimental approaches, designed to avoid parametric assumptions as much as possible, are not well-designed to evaluate the impact on technology choices based on the datasets that are usually available to agricultural production economists (*e.g.*, cost accounting data). Indeed, they either assume conditional independence or rely on specific situations inducing exogenous variations in order to be able to identify all the parameters (*e.g.*, Angrist and Pischke, 2008; Imbens and Wooldridge, 2009; Wooldridge, 2010). Yet, technology choices are production decisions that are similar to input or investment decision. Although they can be impacted by special circumstances, these decisions are much more often made under typical circumstances. Unfortunately, typical circumstances provide limited exogenous variations for investigating the effects of technological choices. This implies that approaches that intend to assess the implications of technology choices should rather be based on parametric models accounting for (i) farmers' technology choices, (ii) the considered input choices and production outcomes and (iii) the links that exist between them.

The standard approaches used by agricultural economists to investigate the impact

of agricultural technology choices at the farm level are presented in the next sections.

1.3 Standard approaches to account for agricultural technology choices

Before presenting the standard approaches to account for agricultural technology choices, let introduce first the CMP – or more generally technology – choice model as

$$\begin{cases} r_{it} &= \mathbb{1}\{m_{it}^* \geq 0\} \\ m_{it}^* &= m(\mathbf{z}_{it}; \boldsymbol{\gamma}) + e_{it} \end{cases}, \quad (1.1)$$

where r_{it} represents the observed CMP choice made by farmer i at time t and m_{it}^* represents the underlying choice process, unobserved to the analyst. In m_{it}^* , e_{it} is an error term, $\boldsymbol{\gamma}$ is a vector of parameters and $m(\cdot)$ denotes a function that is known to the analyst. The equation for m_{it}^* represents the latent regression and measures the net benefit of farmer i using CMP 1 compared to CMP 0 in year t .³ Standard assumption for error term e_{it} is that it has zero mean and follows either a logistic distribution or a normal distribution (see, e.g., Greene, 2020).⁴ From Equation (1.1) and assumptions on e_{it} one can easily derive that

$$\begin{aligned} \text{Prob}[r_{it} = 1 | \mathbf{z}_{it}] &= \text{Prob}[m_{it}^* \geq 0 | \mathbf{z}_{it}] \\ &= \text{Prob}[e_{it} \geq -m(\mathbf{z}_{it}; \boldsymbol{\gamma}) | \mathbf{z}_{it}] \\ &= \text{Prob}[e_{it} < m(\mathbf{z}_{it}; \boldsymbol{\gamma}) | \mathbf{z}_{it}] \\ &= F[m(\mathbf{z}_{it}; \boldsymbol{\gamma})], \end{aligned}$$

where $F(\cdot)$ represents the cumulative distribution function of e_{it} . In particular, we consider here a standard normal distribution with variance equals to 1.

The general objective is to assess the effect of CMP – or more broadly technology – choice r_{it} on (i) y_{it} production function and (ii) $x_{k,it}$ input demands. All the approaches we consider thereafter include model (1.1) for technology choice. They also consider r_{it} – technology choice – as being endogenous regarding y_{it} and \mathbf{x}_{it} – yield and input use demands. We consider two mains cases, depending on how the effect of technology choice on production and input demand functions is modelled. First, in Section 1.3.1, we consider an approach where the technology choice is expected to only create a “shift” in the production and/or input demand functions.

³This benefit includes components that are not purely pecuniary. For instance, farmers’ adopting chemical input saving practices may display above average positive attitudes toward the environment (e.g., Burton, Rigby, and Young, 1999; Mzoughi, 2011) or are valuing different information sources (e.g., Anderson, Jolly, and Green, 2005; Burton, Rigby, and Young, 1999). Indeed, because chemical input saving practices generally require specific agronomic knowledge, information is a key factor in adoption. Technical assistance, extensions services more broadly, are thus to be considered to “boost adoption” (see Piñeiro et al., 2020).

⁴Both distribution are symmetric around their mean, here 0.

Secondly, in Section 1.3.2, we consider an approach where technology choice is considered as a regime, *i.e.* the associated production and input demand functions change based on the implemented regime.

1.3.1 Considering a technology shifting effect in the yield supply and/or input demand models

The first standard approach we consider is when technology r_{it} is introduced as a covariate in the yield supply and input demand models. Such approach was implemented by Burrows (1983) and Fernandez-Cornejo (1996, 1998) when studying the impact of IPM on input uses and/or yields and farm profits. Both authors consider that technology choice only creates a shift in yields and input uses levels, *i.e.* they consider additive technology fixed effects.

The single production choice model with additive technology fixed effects can be described by the following yield supply model:

$$\ln y_{it} = s(\mathbf{w}_{it}; \boldsymbol{\beta}) + \delta r_{it} + v_{it},$$

where v_{it} represents the stochastic part of yield and $s(\cdot)$ corresponds to the production function whose functional form is assumed to be known by the analyst. Price variables and farm(er) characteristics \mathbf{w}_{it} are assumed to be exogenous while r_{it} is expected to be endogenous, *i.e.* $E[v_{it} | \mathbf{w}_{it}, r_{it}] \neq 0$. The standard least squares (LS) estimator of $\boldsymbol{\beta}$ is not consistent because the standard exogeneity assumption does not hold anymore. Yet, we assume that $\mathbf{z}_{it} = (\mathbf{w}'_{it}, \mathbf{q}'_{it})^T$ - covariates from technology choice model – is exogenous in the yield supply model, *i.e.* $E[v_{it} | \mathbf{w}_{it}, \mathbf{q}_{it}] = 0$. This new exogeneity condition allows to build consistent estimators for $(\boldsymbol{\beta}, \delta)$ based on the instrumental variable technique.

Similarly, we can describe the input demand model with additive technology fixed effects with the following equation:

$$\ln x_{k,it} = d_k(\mathbf{w}_{it}; \boldsymbol{\alpha}_k) + \gamma_k r_{it} + u_{k,it},$$

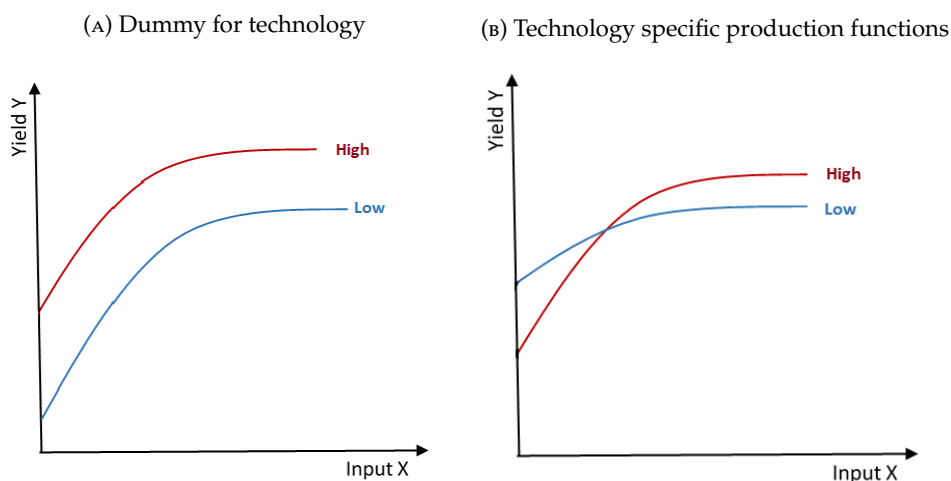
where $u_{k,it}$ represents the stochastic part of input demand and $d_k(\cdot)$ corresponds to the k^{th} input demand function whose functional form is assumed to be known by the analyst. As for the yield supply model, we assume that $E[u_{k,it} | \mathbf{w}_{it}, r_{it}] \neq 0$ but $E[u_{k,it} | \mathbf{w}_{it}, \mathbf{q}_{it}] = 0$ which yields to consistent estimators for parameters $(\boldsymbol{\alpha}_k, \gamma_k)$.

Consistent estimation technique for parameters $(\boldsymbol{\beta}, \delta)$ and $(\boldsymbol{\alpha}_k, \gamma_k)$ include the 2SLS estimation technique as in Fernandez-Cornejo (1996, 1998), or the limited information maximum likelihood (LIML) as in Burrows (1983).⁵

⁵See Greene, 2020, Chapter 8 for more details on these estimation techniques.

By considering technology as a simple covariate, we consider that a change in technology is only shifting yield level y_{it} by δ , the technology parameter (see Figure 1.1 (A)). This assumption of a linear effect of technology choice r_{it} may approximately hold when considering seed varieties or improved input application techniques, at least as long as one ignores heterogeneity in farm production conditions. Yet, it is much more debatable when the considered production technologies significantly differ, as in studies comparing cropping management techniques or irrigated *vs.* non-irrigated crop production. For instance, crop yield response to nitrogen applications depends on seeding dates and densities, or on whether irrigation is available or not, *etc.* As a result, farmers' chemical input uses may display technology specific response to input or crop prices, implying in turn that the functional form of yield supply function $s(\mathbf{w}_{it}; \beta)$ needs to be adapted accordingly (see Figure 1.1 (B)).⁶

FIGURE 1.1: Intercept versus slope change in production functions accounting for technology choices



As a matter of fact, most studies considering the effects of agricultural production technologies assume that crop yield supply and variable input demand functions depend on technologies in a rather flexible way.

⁶These remarks also hold for the input demand model that was not detailed for presentation considerations. The size of the technology shift would correspond to parameter γ_k .

1.3.2 Considering technology as a regime associated to specific production and input demand models

Endogenous regime switching (ERS) model is the standard framework used in studies which intend to assess the impact of crop production technologies adoption on input uses or yield levels. Initially, ERS model was proposed by Lee (1978) for analyzing the effects of unionism on wages. The parametric model this author develops nests two complementary selectivity models as formulated by Heckman (1976, 1978). According to Fuglie and Bosch (1995), the first application of the ERS framework in the agricultural production technology literature comes with Pitt (1983). In the following paragraphs, we present the standard specification and estimation method of ERS models.

Specification of the standard Gaussian ERS model

First, we need to define three sets of latent variables and their respective models. Let y_{it}^r and $\mathbf{x}_{it}^r = (x_{k,it}^r : k \in K)$ respectively define the observed yield level obtained and the chosen input use levels of farmer i in year t given that the farmer uses CMP r . These latent variables are generally modelled as yield supply functions

$$\ln y_{it}^r = s^r(\mathbf{w}_{it}; \boldsymbol{\beta}^r) + v_{it}^r, \quad (1.2)$$

and as sets of input demand functions

$$\ln x_{k,it}^r = d_k^r(\mathbf{w}_{it}; \boldsymbol{\alpha}_k^r) + u_{k,it}^r.$$

If $r_{it} = r$, $r \in \{0, 1\}$, potential yield and input uses levels are observed with $(y_{it}^r, \mathbf{x}_{it}^r) = (y_{it}, \mathbf{x}_{it})$. Otherwise, they are counterfactual. Indeed, in this specification $(y_{it}, \mathbf{x}_{it})$ are latent variables that are linked to their observed counterparts through the following regime switching equations

$$\begin{cases} y_{it} & = r_{it}y_{it}^1 + (1 - r_{it})y_{it}^0 \\ x_{k,it} & = r_{it}x_{k,it}^1 + (1 - r_{it})x_{k,it}^0 \end{cases}. \quad (1.3)$$

The standard ERS model is also completed by the CMP choice model defined in Equation (1.1) by

$$\begin{cases} r_{it} & = \mathbb{1}\{m_{it}^* \geq 0\} \\ m_{it}^* & = m(\mathbf{z}_{it}; \boldsymbol{\gamma}) + e_{it} \end{cases}.$$

CMP choice r_{it} is endogenous regarding yield supply level when error terms e_{it} and (v_{it}^0, v_{it}^1) are correlated. Before deriving the consequences of such correlation, let consider a simple Roy model to illustrate why correlated error terms is likely to be the rule rather than the exception when assessing the effects of technology choices

on production choices and outcomes. The assumptions underlying the so-called Roy model state that the latent index underlying the CMP choice model is simply given by $m_{it} = \ln y_{it}^1 - \ln y_{it}^0$ (see, e.g., Heckman and Honore, 1990; Suri, 2011). According to this model, farmers choose the CMP associated to the highest yield level.⁷ It implies that error term $e_{it} = v_{it}^1 - v_{it}^0$ is positively correlated with v_{it}^1 and/or negatively correlated with v_{it}^0 . Such correlations indicate the presence of the so-called self-selection effects. When choosing their CMP, farmers compare the potential outcomes of the CMPs and to do this comparison they account for factors and conditions that are not observed by the analyst. Despite the restrictiveness of Roy model assumptions, farmers' expectations of yield level ratios y_{it}^1/y_{it}^0 can be expected to be major drivers of CMP choices r_{it} and these expectations certainly depend on factors that are unknown to the analyst. Such remarks also hold for input uses models and confirm our self-selection hypothesis.

Other than self-selection effects are selection biases. Self-selection relies on causal effects whereas selection biases occur due to confounding mechanism. It means that both error terms e_{it} and v_{it}^r (and respectively $u_{k,it}^r$) contain the effects of common unobserved factors. In the case of CMP adoption, one can think that it is easier to implement new technologies for highly skilled farmers. Their skills can also impact their input uses or yield levels. If not observed and not accounted for, farmers' skills yield to biased estimates. More generally, unobserved factors affecting both farmers' technology and production choices lead to selection biases if not accounted for. The presence of such potential selection biases speaks in favor of error term correlation in the agricultural production technology framework.

When e_{it} and (v_{it}^1, v_{it}^0) are correlated, parameter β^r cannot be estimated only based on the sub-sample of farmers using CMP r without accounting for their choosing to use this CMP. In particular, from e_{it} and v_{it}^r correlation derives the fact that $E[v_{it}^r | \mathbf{z}_{it}, r_{it} = r] \neq 0$, even if exogeneity condition $E[v_{it}^r | \mathbf{z}_{it}] = 0$ holds. Omitting this observation and estimating β^r conditionally on $r_{it} = r$ (i.e., based on the sub-sample of farmers using CMP r) using ordinary least squares (LS) in Model (1.2) would result in biased estimates of β^r . Indeed, we can write the conditional expectation of y_{it} on \mathbf{z}_{it} and $r_{it} = r$ as

$$E[\ln y_{it}^r | \mathbf{z}_{it}, r_{it} = r] = s^r(\mathbf{w}_{it}; \beta^r) + E[v_{it}^r | \mathbf{z}_{it}, r_{it} = r],$$

where $E[v_{it}^r | \mathbf{z}_{it}, r_{it} = r]$ is a non-trivial function of \mathbf{z}_{it} that generates an estimation bias when it is ignored. As for input use models and their estimation, the reasoning

⁷According to this schematic model farmers ignore input costs, which is basically harmless when these don't depend much on CMPs. Farmers are also assumed to perfectly anticipate the technology specific production levels. This assumption does not hold for crop production, which is highly random to its dependence on weather effects, on weed competition and on pest and diseases outbreaks.

still applies *i.e.*

$$E[\ln x_{k,it}^r | \mathbf{z}_{it}, r_{it} = r] = d_k^r(\mathbf{w}_{it}; \boldsymbol{\alpha}_k^r) + E[u_{k,it}^r | \mathbf{z}_{it}, r_{it} = r].$$

Non parametrically estimating conditional expectation terms $E[v_{it}^r | \mathbf{z}_{it}, r_{it} = r]$ and $E[u_{k,it}^r | \mathbf{z}_{it}, r_{it} = r]$ is possible but difficult. Furthermore, this requires the information content of the technology choice instrumental variables, \mathbf{q}_{it} , to be sufficient, which is rarely the case. Yet, ERS models generally make the most of their fully parametric form. Self-selection effects and selection biases are accounted for by explicitly modelling the error term correlations. For estimation purposes, pairs (e_{it}, v_{it}^r) and $(e_{it}, u_{k,it}^r)$ are typically assumed to be normal and not to depend on \mathbf{z}_{it} for $r \in \{0, 1\}$, *i.e.*:

$$(v_{it}^r, e_{it}) | \mathbf{z}_{it} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Omega}_{ve}^r) \text{ with } \boldsymbol{\Omega}_{ve}^r = \begin{bmatrix} \omega_{vv}^r & \omega_{ve}^r \\ \omega_{ve}^r & \omega_{ee}^r \end{bmatrix},$$

and

$$(u_{k,it}^r, e_{it}) | \mathbf{z}_{it} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Omega}_{k,ue}^r) \text{ with } \boldsymbol{\Omega}_{k,ue}^r = \begin{bmatrix} \omega_{k,uu}^r & \omega_{k,ue}^r \\ \omega_{k,ue}^r & \omega_{ee}^r \end{bmatrix}.$$

Under these assumptions, the considered model is fully parametric and parameters $(\boldsymbol{\beta}^0, \boldsymbol{\beta}^1, \boldsymbol{\gamma})$ or $(\boldsymbol{\alpha}_k^0, \boldsymbol{\alpha}_k^1, \boldsymbol{\gamma})$ can be estimated by using Maximum Likelihood (ML) or Heckman's well known two-step estimators (*e.g.*, Heckman, Tobias, and Vytlačil, 2003; Wooldridge, 2002). In particular, under these assumptions technology choices follow a standard Probit model conditionally on \mathbf{z}_{it} , implying that parameter $\boldsymbol{\gamma}$ can easily be estimated in a first step. More importantly, these joint normality assumptions allow error term e_{it} and error terms $(v_{it}^r, u_{k,it}^r)$ to be correlated. Yet, these normality assumptions impose these links to be linear. These links are fully described by fixed covariance parameters ω_{ve}^r and $\omega_{k,ue}^r$.

Estimation of the standard Gaussian ERS model

Heckman's two-step estimator makes use of well-known results on the effects of auxiliary truncations on the moments of jointly normal variables. These results allow to devise control functions to account for endogenous sample selection issues. Those control functions are then used in the second step of Heckman's estimation approach. From the results on the effects of auxiliary truncations we can derive the following means of the error terms of the yield and input use level models conditional on technology choices:

$$E[v_{it}^r | \mathbf{z}_{it}, r_{it} = r] = (\omega_{ee})^{-1/2} \omega_{ve}^r \lambda^r \left((\omega_{ee})^{-1/2} m(\mathbf{z}_{it}; \boldsymbol{\gamma}) \right), \quad (1.4)$$

and

$$E[u_{k,it}^r | \mathbf{z}_{it}, r_{it} = r] = (\omega_{ee})^{-1/2} \omega_{k,ue}^r \lambda^r \left((\omega_{ee})^{-1/2} m(\mathbf{z}_{it}; \boldsymbol{\gamma}) \right), \quad (1.5)$$

where $\lambda(\cdot)$ denotes the inverse Mills ratio (IMR). The IMR corresponds to:

$$\lambda(a) = \Phi(a)^{-1} \varphi(a),$$

where function $\varphi(\cdot)$ denotes the probability distribution function of $\mathcal{N}(0, 1)$ and $\Phi(\cdot)$ its cumulative distribution function. Moreover, function $\lambda^r(\cdot)$ is defined by

$$\lambda^r(a) = (2r - 1)\lambda((2r - 1)a).$$

If $\omega_{ve}^r = 0$ and $\omega_{k,ue}^r = 0$, then error term e_{it} is uncorrelated with error terms v_{it}^r and $u_{k,it}^r$. In that particular case, technology choice endogeneity can be ignored conditionally on \mathbf{z}_{it} in the yield and input use levels models. Otherwise, when such exogeneity conditions are not met, Equations (1.4) and (1.5) provide control functions for the second step of Heckman's estimation approach. Thus, we can define the following "augmented" regression equations for estimating the technology specific yield supply models:

$$\ln y_{it}^r = s^r(\mathbf{w}_{it}; \boldsymbol{\beta}^r) + \omega_{ve}^r \lambda^r(m(\mathbf{z}_{it}; \boldsymbol{\gamma})) + \mu_{k,it}^r \quad (1.6)$$

and input demand models:

$$\ln x_{k,it}^r = d_k^r(\mathbf{w}_{it}; \boldsymbol{\alpha}_k^r) + \omega_{k,ue}^r \lambda^r(m(\mathbf{z}_{it}; \boldsymbol{\gamma})) + \eta_{k,it}^r \quad (1.7)$$

with $E[\mu_{k,it}^r | \mathbf{z}_{it}, r_{it} = r] = 0$ and $E[\eta_{k,it}^r | \mathbf{z}_{it}, r_{it} = r] = 0$. Generally, the normalization condition $\omega_{ee} = 1$ is imposed, as it is the case in Equations (1.6) and (1.7).

Before estimating such "augmented models", we need to estimate parameter $\boldsymbol{\gamma}$ to compute an estimate of Mills ratio terms $\lambda^r(m(\mathbf{w}_{it}, \mathbf{q}_{it}; \boldsymbol{\gamma}))$. The estimation of parameter $\boldsymbol{\gamma}$ corresponds to the first step of Heckman's two step estimation procedure. Then, LS estimators based on the farm sub-sample actually using CMP r can be used for consistently estimating parameter of the "augmented" models *i.e.* $(\boldsymbol{\beta}^r, \omega_{ve}^r)$ and $(\boldsymbol{\alpha}_k^r, \omega_{k,ue}^r)$.⁸ When functions $s^r(\mathbf{w}_{it}; \boldsymbol{\beta}^r)$ and $d_k^r(\mathbf{w}_{it}; \boldsymbol{\alpha}_k^r)$ are linear in their parameters, linear LS can be implemented.

The estimation of ERS models relies on instrumental variables \mathbf{q}_{it} . Effectively, these variables are part of the CMP choice drivers \mathbf{z}_{it} but are excluded from the drivers of the input demand and yield supply functions \mathbf{w}_{it} . Thus, inducing exogenous variations in Mills ratio terms $\lambda^r(m(\mathbf{w}_{it}, \mathbf{q}_{it}; \boldsymbol{\gamma}))$ contributes to separately identify parameters $\boldsymbol{\beta}^r$ and ω_{ve}^r in the yield supply equation (Equation (1.6)). In the absence of such instrumental variable vector, the separate identification of parameters

⁸From that second step we can also derive a test for the exogeneity of technology choices conditionally on \mathbf{z}_{it} . Indeed, null hypotheses $\omega_{ve}^r = 0$ or $\omega_{k,ue}^r = 0$ that are tested with the usual Student or Chi-2 tests, correspond to the absence of technology endogeneity. Given that under the null hypotheses, the first step of the estimation process does not affect the second step estimators, the standard tests can be used without correcting for multistage estimation.

β^r and ω_{ve}^r only relies on the functional form differences that distinguish the effects of \mathbf{w}_{it} in structural function $s^r(\mathbf{w}_{it}; \beta^r)$ and in control function $\omega_{ve}^r \lambda^r(m(\mathbf{z}_{it}; \gamma))$. Such functional form identification strategy is unwarranted as it relies on questionable, and partly untestable, distributional assumptions (e.g., Wooldridge, 2015). Yet, it is often difficult to find instrumental variables with sufficient information content to ensure the identification of the model parameters in a non-parametric sense. Admittedly, factors impacting technology choices are likely to impact input choices and production outcomes as well. Thus, identification of the model parameters generally relies on both the effects of instrumental variable \mathbf{q}_{it} in control functions $\omega_{ve}^r \lambda^r(m(\mathbf{z}_{it}; \gamma))$ and $\omega_{k,ue}^r \lambda^r(m(\mathbf{z}_{it}; \gamma))$, as well as on the functional form differences distinguishing the effects of control variables \mathbf{w}_{it} in the control functions and in the core of the model, *i.e.* in functions $s^r(\mathbf{w}_{it}; \beta^r)$ and $d_k^r(\mathbf{w}_{it}; \alpha_k^r)$.

Despite being easy to implement, Heckman's two-step estimator is less efficient than its ML counterpart. This is due to the fact that it fails to account for statistical element linking the estimation criterion of γ and those of the other parameters. Additionally, even if it is relatively easy, determining the asymptotic distributions of the second step estimators, which depend on that of the estimator of γ , remains a tedious task. Resampling techniques can be used for estimating the empirical distribution of the second step estimators. Yet, even if easy to implement, these techniques are computationally intensive.

Some extensions of the standard ERS model

The benefit of ERS models as presented above is that they allow to estimate average causal effect of technology choices. In particular, we can derive from the ERS model the average causal effect of CMP 1 (compared to CMP 0) on yield level, *i.e.* $E[\ln y_{it}^1 - \ln y_{it}^0]$. One can also estimate the average causal effect of CMP 1 for farmer using CMP r , *i.e.* $E[\ln y_{it}^1 - \ln y_{it}^0 | r_{it} = r]$.⁹ Overall, ERS models permit to estimate the overall impact of technology choice on input uses or yield levels but also the differentiated impact for each farm type. Up to our knowledge, Di Falco, Veronesi, and Yesuf (2011) present the first application of this approach comparing the effects of production technologies on agricultural production choices or outcomes. Not to be forgotten are the restrictive parametric and distributional assumptions that allow to draw such average causal effects of "treatments". Yet, ERS is among the few approaches that are applicable when the information content of the control variables does not permit to correct entirely the self-selection effects and/or selection biases.

⁹In a context of endogenous covariates, usual formulas for estimating the causal effect of CMP choice need to be adjusted for.

As originally developed by Lee (1978), the ERS model is designed to compare the effects of two technologies: the conventional and the innovative alternative technology (see, *e.g.*, Abdulai and Huffman, 2014; Di Falco, Veronesi, and Yesuf, 2011; Fuglie and Bosch, 1995; Kabunga, Dubois, and Qaim, 2012; Khanal et al., 2018b; Pitt, 1983; Shiferaw et al., 2014, for applications to agricultural production technology). Some extensions of the basic ERS model were proposed by (i) Wu and Babcock (1998) and Di Falco and Veronesi (2013) for more than two technologies and by (ii) Khanna (2001) for sequential adoption of several technologies. Next section is dedicated to the presentation of the extension proposed by Suri (2011), an extension requiring panel data.

1.3.3 Technology specific production function in a panel data ERS model

The extension of Suri (2011) makes the most of panel data by developing an ERS model free of (parametric) technology choice model. The resulting model is based on technology specific yield models given by

$$\ln y_{it}^r = \beta_t^r + \ln \mathbf{x}'_{it} \boldsymbol{\beta}_x^r + v_{it}^r,$$

where $v_{it}^r = v_i^r + \mu_{it}^r$. The error term v_{it}^r is defined as the sum of a farm and technology specific random term, v_i^r , and of a technology specific idiosyncratic error term, μ_{it}^r . Error term v_i^r can also be decomposed into the sum of two independent terms, $v_i^r = \gamma^r \theta_i + \tau_i$ (*e.g.*, Heckman and Honore, 1990; Lemieux, 1998). $\gamma^r \theta_i$ represents the technology specific term and captures the comparative advantage of technology 1 over technology 0. On the other hand, τ_i is a farm(er) specific term capturing the absolute productivity advantage. This decomposition is finally used for transforming regime switching equation $\ln y_{it} = r_{it} \ln y_{it}^1 + (1 - r_{it}) \ln y_{it}^0$ into a “generalized” yield model.

The resulting yield model is a correlated random coefficient model that can be estimated using an extension of the linear projection approach proposed by Chamberlain (1982, 1984). Yet, for the parameters to be correctly identified, technology choice r_{it} needs to be uncorrelated with error terms μ_{it}^r and farm specific parameter τ_i (although r_{it} can, and is likely to, be correlated with comparative advantage random term θ_i). The most appealing characteristic of this ERS modelling framework is the fact that it does not rely on restrictive functional form assumptions. In particular, technology choice does not need to be modelled. The counterpart of that is the need for (i) observed technology choices for several consecutive years for most farmers and (ii) sufficient technology changes in the dataset.¹⁰

¹⁰Also, identification of the considered model rests on unusual assumptions regarding input uses, \mathbf{x}_{it} . First, input uses are assumed to not depend on technology choice r_{it} . This assumption is unusual for ERS models considering agricultural production technology choices. Second, input uses are assumed exogenous with respect to error terms μ_{it}^r and, especially, to farm specific random term τ_i . The strict exogeneity assumption stating that $E[\mu_{it}^r | \mathbf{x}_{it}] = 0$ is common in panel data models of production

1.4 A new modelling framework to account for technology choice in agricultural production using the primal

Despite the wide variety of approaches proposed in the literature – most standard ones we presented erstwhile – for comparing implications of technology choice on crop production choices and outcomes, any of these approaches fully fulfill our needs. As a matter of fact, we intend to assess the effects of, low-input *versus* high-input, CMPs on the production function describing how wheat yield responds to chemical input use levels. From that derives the fact that we need to consider primal yield functions instead of dual yield supply functions. Yet, when measuring the effects of production technologies on farmers' choices and production outcomes, most studies consider input demand and yield supply models. The studies of Suri (2011) and Michler and Shively (2015) are notable exceptions in this respect.

Considering primal yield functions implies that the input endogeneity issues need to be addressed to get consistent estimates of the production function (*e.g.*, Akerberg, Caves, and Frazer, 2015; Griliches and Mairesse, 1995).¹¹ Such endogeneity issues are well-known. For instance, farmers adapt their fertilizer and pesticide uses to the conditions prevailing in their fields. While they are observed by farmers, these conditions are generally poorly documented in standard datasets, implying that their effects are captured by the random terms of the yield functions and that these random terms are likely to be correlated with chemical input uses. As pointed out by Akerberg, Caves, and Frazer (2015), applied economists have devoted much attention to this problem and to approaches to solve it.¹² Yet, to our knowledge, none of those approaches were considered when estimating production functions in a ERS framework. Indeed, estimating technology specific production functions in an ERS model requires us to simultaneously address the endogeneity issue of input uses and the one related to the technology choice.

The low-input CMP is primarily intended to reduce pesticide use – a shared objective for most public authorities. Thus, we pay special attention to the effects of these particular inputs in the yield functions we consider. Following the agricultural production literature, and in particular the work from Lichtenberg and Zilberman (1986), we consider a yield functions in which pesticides impact crop production

functions. Assuming that input uses and the farm specific effect is much less frequent, due to the seminal work of Mundlak (1961) in particular (see, *e.g.*, Griliches and Mairesse, 1995).

¹¹Input use endogeneity in agricultural production functions is often used as a typical example in the econometric literature considering endogeneity issues in general (*e.g.*, Chamberlain, 1982, 1984; Mundlak, 1961, 1978, 1996; Zellner, Kmenta, and Dreze, 1966).

¹²See Akerberg, Caves, and Frazer (2015) for a brief overview of the related literature and Greene, 2020, Chapter 8.

through their effects in a damage abatement function.¹³ The use of damage abatement function highlights the protective role of pesticides on crop yields whereas other inputs are generally considered as purely productive inputs that are increasing the potential yield level. Yet, yield functions featuring damage abatement functions are generally nonlinear in their parameters. Instrumental variables techniques, usually considered for dealing with input use endogeneity, are more complex to use in a nonlinear framework. In particular, standard nonlinear 2SLS estimators often perform poorly in practice. Solutions to this problem have been proposed, including estimation approaches based on Generalized Method of Moments (GMM) estimators using estimated (near-)efficient instruments, but their practical implementation is not straightforward. The estimation approach that we propose makes use of instrumental variables to determine control functions instead of estimating orthogonality conditions (*e.g.*, Wooldridge, 2015).¹⁴

Our estimation framework can be seen as an extension to Heckman’s two-step approach for estimating regression models under endogenous sample selection with endogenous regressors. We use a control function that accounts for the correlation of the endogenous regressors and we adapt the expression of the Mills ratio that is used for controlling for endogenous sample selection in the approach of Heckman. Another extension of Heckman’s two-step estimation approach under endogenous sample selection was proposed by Wooldridge (2010) for handling the issues raised by endogenous regressors. This approach relies on the usual Mills ratios to deal with endogenous sample selection but uses 2SLS estimators, with Mills ratio terms included in the instrument vectors, for accounting for regressor endogeneity. This approach is well suited for linear models but requires significant adaptations for estimating models with nonlinear functional forms. The approach we developed here is easier to implement in practice than its competing alternatives, although it requires distributional assumptions that are more restrictive.

1.4.1 Technology specific yield and input demand functions in a ERS framework

To investigate the effects of the low-input *versus* high-input CMP choice on wheat yield response to chemical input uses, we consider an ERS model that combines (*i*) a pair of CMP specific yield function and input demand systems:

$$(y_{it}, \mathbf{x}_{it}) = (y_{it}^r, \mathbf{x}_{it}^r) \text{ if } r_{it} = r, \quad (1.8)$$

¹³See Frisvold (2019) for a literature survey on the measurement of the effects of pesticides on crop production.

¹⁴Kutlu (2010) proposed a similar approach, which has been widely used since then (*e.g.*, Amsler, Prokhorov, and Schmidt, 2016; Kumbhakar, Parmeter, and Zelenyuk, 2020), for dealing with the same issues in stochastic production frontier models.

where $\begin{cases} \ln y_{it}^r &= \ln f^r(\mathbf{x}_{it}^r, \mathbf{c}_{it}; \boldsymbol{\beta}^r) + v_{it}^r \\ \ln x_{k,it}^r &= \alpha_{k,0}^r + \mathbf{w}'_{it} \boldsymbol{\alpha}_{k,w}^r + u_{k,it}^r \end{cases} \quad k \in K, \text{ with (ii) a standard CMP choice model:}$

$$r_{it} = \mathbb{1}\{\gamma_0 + \boldsymbol{\gamma}'_z \mathbf{z}_{it} + e_{it} \geq 0\}. \quad (1.9)$$

As evoked before, the functional form of CMP specific yield functions $f^r(\cdot; \boldsymbol{\beta}^r)$ is nonlinear in parameter vector $\boldsymbol{\beta}^r$. Input demand models and the index of the CMP choice are assumed to be linear in their explanatory variables and parameters.¹⁵ Additional assumption is that all considered models entail additively separable error terms (respectively, v_{it}^r , \mathbf{u}_{it}^r and e_{it} , with $\mathbf{u}_{it}^r = (u_{k,it}^r : k \in K)$) as their unique random terms. In particular, it means that farmers' chemical input choices respond homogeneously to price changes.¹⁶

Unlike other ERS models that have been considered in the literature, our model considers endogenous regressors. Specifically, we address the potential correlation between \mathbf{x}_{it}^r and v_{it}^r that comes from the fact that chemical input use levels \mathbf{x}_{it}^r are likely to be endogenous in the corresponding yield functions, with $\mathbf{x}_{it}^r = (x_{k,it}^r : k \in K)$ the vector of input uses. We assume that variables \mathbf{z}_{it} are exogenous with respect to error terms v_{it}^r and \mathbf{u}_{it}^r . From that we derive that \mathbf{x}_{it}^r and v_{it}^r are correlated conditional on \mathbf{z}_{it} only if error terms v_{it}^r and \mathbf{u}_{it}^r are correlated. We assume that error term vectors $(v_{it}^r, \mathbf{u}_{it}^r, e_{it})$ are jointly normal and independent of control and instrumental variables \mathbf{z}_{it} , with

$$(\mathbf{u}_{it}^r, v_{it}^r, e_{it}) | \mathbf{z}_{it} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Omega}^r) \text{ with } \boldsymbol{\Omega}^r = \begin{bmatrix} \boldsymbol{\Omega}_{uu}^r & \boldsymbol{\Omega}_{uv}^r & \boldsymbol{\Omega}_{ue}^r \\ (\boldsymbol{\Omega}_{uv}^r)' & \omega_{vv}^r & \omega_{ve}^r \\ (\boldsymbol{\Omega}_{ue}^r) & \omega_{ve}^r & 1 \end{bmatrix} \text{ for } r \in \{0, 1\}, \quad (1.10)$$

where

$$\boldsymbol{\Omega}_{uv}^r = (\omega_{k,uv}^r : k \in K), \boldsymbol{\Omega}_{ue}^r = (\omega_{k,ue}^r : k \in K) \text{ and } \boldsymbol{\Omega}_{uu}^r = [\omega_{k\ell,uu}^r : (k, \ell) \in K \times K].$$

Apart from being usually imposed, the well-known properties of multivariate normal variables make such assumption convenient to deal with the multiple endogeneity issues we face.

1.4.2 Insights for the estimation procedure

Despite our considering a fully parametric ERS model, the nonlinear production function and the numerous parameters to be estimated advocate for an alternative estimation procedure to the standard ML. The distributional assumptions given in Equation (1.10) enable us to define an alternative, multistage estimation approach for

¹⁵These functional forms are obviously convenient for estimation purpose but more general functional forms could be considered without significantly impacting the estimation process of the model.

¹⁶Koutchadé, Carpentier, and Féménia (2018, 2020) recently provided empirical evidences showing that farmers respond heterogeneously to economic incentives.

our ERS model. This estimation approach is fairly simple as it consists in a sequence of estimation problems that are easy to solve, *i.e.* Probit model and LS estimation problems. The estimation approach we present thereafter can be seen as an extension of Heckman's two-step approach for estimating standard – *i.e.* with exogenous explanatory variables – Gaussian ERS models (*e.g.*, Heckman, Tobias, and Vytlačil, 2003; Wooldridge, 2002, 2015). It relies on two sets of control functions. The first one is used to deal with the input use endogeneity issue in the production function while the second one is used to deal with the sample selection issues due to the production practice choice.

First, we proceed with the estimation of the CMP specific input use models. One can define a standard, linear and Gaussian, ERS model for each considered input with the following input demand model:

$$\ln x_{k,it}^r = \alpha_{k,0}^r + \mathbf{w}'_{it} \boldsymbol{\alpha}_{k,w}^r + u_{k,it}^r \text{ when } r_{it} = r. \quad (1.11)$$

Then, we make use of Heckman's standard result (Heckman, 1976, 1979):

$$E[u_{k,it}^r | \mathbf{z}_{it}, r_{it} = r] = \omega_{k,ue}^r \lambda^r (\gamma_0 + \boldsymbol{\gamma}'_z \mathbf{z}_{it}),$$

to get augmented regression equations given by:

$$\ln x_{k,it} = \alpha_{k,0}^r + \mathbf{w}'_{it} \boldsymbol{\alpha}_{k,w}^r + \omega_{k,ue}^r \lambda^r (\gamma_0 + \boldsymbol{\gamma}'_z \mathbf{z}_{it}) + \eta_{k,it}^r, \quad (1.12)$$

with $E[\eta_{k,it}^r | \mathbf{z}_{it}, r_{it} = r] = 0$. Equation (1.12) defines regression models in the sub-samples defined by the farming system choices. Term $\omega_{k,ue}^r \lambda^r (\gamma_0 + \boldsymbol{\gamma}'_z \mathbf{z}_{it})$ defines a control function for endogenous selection of the observations characterized by $r_{it} = r$ in the considered sub-sample. Error term $\eta_{k,it}^r$ is defined by $\eta_{k,it}^r = u_{k,it}^r - \omega_{k,ue}^r \lambda^r (\gamma_0 + \boldsymbol{\gamma}'_z \mathbf{z}_{it})$, which implies that exogeneity condition $E[\eta_{k,it}^r | \mathbf{z}_{it}, r_{it} = r] = 0$ necessarily holds (in the sub-sample characterized by $r_{it} = r$).

Input demand functions can thus be estimated following Heckman's two-step approach. First, the Probit model of $r_{it} | \mathbf{z}_{it}$ (defined by Equation (1.9)) is estimated by ML for obtaining a consistent estimate of parameter $\boldsymbol{\gamma}$, $\hat{\boldsymbol{\gamma}} = (\hat{\gamma}_0, \hat{\boldsymbol{\gamma}}_z)$. This estimate can then be used for obtaining consistent estimates of Mills ratio terms, $\lambda^r (\hat{\gamma}_0 + \hat{\boldsymbol{\gamma}}_z \mathbf{z}_{it})$. Second, consistent estimates of parameters $(\alpha_{k,0}^r, \boldsymbol{\alpha}_{k,w}^r, \omega_{k,ue}^r)$ can be obtained by applying standard linear LS to Equation (1.12) in the sub-sample characterized by $r_{it} = r$ after replacing terms $\lambda^r (\gamma_0 + \boldsymbol{\gamma}'_z \mathbf{z}_{it})$ by their estimates. Let vectors $(\hat{\alpha}_{k,0}^r, \hat{\boldsymbol{\alpha}}_{k,w}^r, \hat{\omega}_{k,ue}^r)$ denote these parameter estimates.

Let now consider the following CMP specific yield function:

$$\ln y_{it} = \ln f^r(\mathbf{x}_{it}, \mathbf{c}_{it}; \boldsymbol{\beta}^r) + v_{it}^r.$$

From the joint distribution of error term vectors $(v_{it}^r, \mathbf{u}_{it}^r, e_{it})$ (cf. Equation 1.10) and technology choice model, we can write:¹⁷

$$E[\ln y_{it} | \mathbf{x}_{it}, \mathbf{z}_{it}, \mathbf{u}_{it}^r, r_{it} = r] = \ln f^r(\mathbf{x}_{it}, \mathbf{c}_{it}; \boldsymbol{\beta}^r) + (\boldsymbol{\rho}_{vu}^r)' \mathbf{u}_{it}^r + \psi_{ve}^r \lambda_{it}^r, \quad (1.13)$$

which is the first key result for our estimation approach, where:

$$\lambda_{it}^r = (\psi_{ee}^r)^{-1/2} \lambda^r \left((\psi_{ee}^r)^{-1/2} (\gamma_0 + \boldsymbol{\gamma}'_z \mathbf{z}_{it} + (\boldsymbol{\rho}_{eu}^r)' \mathbf{u}_{it}^r) \right), \quad (1.14)$$

and

$$\begin{aligned} \boldsymbol{\rho}_{vu}^r &= (\boldsymbol{\Omega}_{uu}^r)^{-1} \boldsymbol{\omega}_{uv}^r, \\ \boldsymbol{\rho}_{eu}^r &= (\boldsymbol{\Omega}_{uu}^r)^{-1} \boldsymbol{\omega}_{ue}^r, \\ \psi_{ve}^r &= \omega_{ve}^r - (\boldsymbol{\omega}_{uv}^r)' (\boldsymbol{\Omega}_{uu}^r)^{-1} \boldsymbol{\omega}_{ue}^r, \\ \psi_{ee}^r &= 1 - (\boldsymbol{\omega}_{ue}^r)' (\boldsymbol{\Omega}_{uu}^r)^{-1} \boldsymbol{\omega}_{ue}^r. \end{aligned} \quad (1.15)$$

In Equation (1.13), term $(\boldsymbol{\rho}_{vu}^r)' \mathbf{u}_{it}^r$ is a control function for the endogeneity of $\mathbf{x}_{it} = \mathbf{x}_{it}^r$ while term $\psi_{ve}^r \lambda_{it}^r$ is a control function for the selection of observations characterized by $r_{it} = r$. Error term \mathbf{u}_{it}^r can be defined as a function of exogenous variables that include price vector \mathbf{p}_{it} . This vector plays the role of instrumental variables of the endogenous input use levels \mathbf{x}_{it}^r in the model of $\ln y_{it}^r$. Equation (1.14) shows adding the usual Mills ratio, $\lambda^r(\gamma_0 + \boldsymbol{\gamma}'_z \mathbf{z}_{it})$, is inappropriate for controlling for sample selection in the yield model with endogenous input uses. Term λ_{it}^r has error term \mathbf{u}_{it}^r as an argument and needs to be scaled by parameter $(\psi_{ee}^r)^{-1/2}$. Equations (1.13) and (1.14) imply that a model of $\ln y_{it}$ can be written as

$$\ln y_{it} = \ln f^r(\mathbf{x}_{it}, \mathbf{c}_{it}; \boldsymbol{\beta}^r) + (\boldsymbol{\rho}_{vu}^r)' \mathbf{u}_{it}^r + \psi_{ve}^r \lambda_{it}^r + \mu_{it}^r, \quad (1.16)$$

with $E[\mu_{it}^r | \mathbf{x}_{it}, \mathbf{z}_{it}, \mathbf{u}_{it}^r, \lambda_{it}^r, r_{it} = r] = 0$. Error term μ_{it}^r is defined by $\mu_{it}^r = v_{it}^r - (\boldsymbol{\rho}_{vu}^r)' \mathbf{u}_{it}^r - \psi_{ve}^r \lambda_{it}^r$. Equation (1.16) states that $\ln y_{it}$ follows a regression model in $(\mathbf{x}_{it}, \mathbf{c}_{it}, \mathbf{u}_{it}^r, \lambda_{it}^r)$ conditionally on $r_{it} = r$. Accordingly, consistent estimates of parameters $\boldsymbol{\beta}^r$, $\boldsymbol{\rho}_{vu}^r$ and ψ_{ve}^r can be obtained by applying non linear LS to the “augmented” regression model given in Equation (1.16) based on the sub-sample of farmer using farming system r . Yet, implementing this approach supposes that error terms \mathbf{u}_{it}^r and Mills ratio terms λ_{it}^r can be estimated *a priori*. Consistent estimates of error terms vectors \mathbf{u}_{it}^r , denoted by $\hat{\mathbf{u}}_{it}^r = (\hat{u}_{k,it}^r : k \in K)$, are easily obtained by using the residual terms of the input demand models, that is to say terms $u_{k,it}^r = \ln x_{k,it}^r - \alpha_{k,0}^r - \mathbf{w}'_{it} \boldsymbol{\alpha}_{k,w}^r$ for $k \in K$ and $r \in \{0, 1\}$.

The second key result on which our estimation is based implies that parameters ψ_{ee}^r

¹⁷Calculation details leading to this first key result can be found in Appendix 1.6.1.

and ρ_{eu}^r can be estimated *a priori*, implying that terms λ_{it}^r can also be estimated *a priori*. Indeed, given that we can consistently estimate ω_{ue}^r *a priori*, definitions of ψ_{ee}^r and ρ_{eu}^r (cf. Equation (1.15)) show that we just need to obtain a consistent estimate of Ω_{uu}^r – i.e. parameters $\omega_{k\ell,uu}^r$ for $k \in K$ – for obtaining consistent estimates of parameters ψ_{ee}^r and ρ_{eu}^r , and thus of terms λ_{it}^r . Terms $\omega_{k\ell,uu}^r$ are marginal covariance parameters. Howbeit, we can only get estimates of error terms $u_{k,it}^r$ when $r_{it} = r$. Yet, we can easily demonstrate that, to obtain $\omega_{k\ell,uu}^r$, we just need to correct $E[u_{k,it}^r u_{\ell,it}^r | r_{it} = r]$ for yielding $E[u_{k,it}^r u_{\ell,it}^r]$.¹⁸ This result is given by:

$$\omega_{k\ell,uu}^r = E[u_{k,it}^r u_{\ell,it}^r | r_{it} = r] + \omega_{k,ue}^r \omega_{\ell,ue}^r E[(\gamma_0 + \gamma'_z \mathbf{z}_{it}) \lambda^r (\gamma_0 + \gamma'_z \mathbf{z}_{it}) | r_{it} = r], \quad (1.17)$$

for $(k, \ell) \in K^2$. All parameters involved in the right hand side term of Equation (1.17) can be estimated *a priori*. Conditional expectation $E[u_{k,it}^r u_{\ell,it}^r | r_{it} = r]$ is consistently estimated in the input demand model by its empirical counterpart in the sub-sample characterized by $r_{it} = r$. Consistent estimates of parameters $\omega_{k,ue}^r$ and (γ_0, γ_z) are also obtained when estimating respectively the input demand and technology choice models. As a result, the sample counterpart of the right hand side term of Equation (1.17) provides a consistent estimator of parameter $\omega_{k\ell,uu}^r$, implying that a consistent estimate of matrix Ω_{uu}^r is easily obtained.

The above key results – which are fully described in Appendices 1.6.1 and 1.6.1 - and insights yield an estimation procedure for estimating an ERS model with endogenous regressors. To summarize this procedure consists of three main steps. First, we estimate the technology choice model. Results of this estimation step provide elements for computing control functions that are used in the next step. Second, we estimate the ERS input use models. Once again, results obtained in this estimation step are used for computing control functions to be used in the next step. The ERS yield function models are estimated in the last step. Empirical estimates of the resulting multistep estimator can be obtained based on resampling techniques. The detailed estimation procedure is detailed thereafter:

(A.1) Compute the ML estimate of (γ_0, γ_z) , $(\hat{\gamma}_0, \hat{\gamma}_z)$, by estimating the Probit model of $r_{it} | \mathbf{z}_{it}$ based on the full sample.

(A.2) Compute the estimates of the Mills ratio terms $\lambda^r(\gamma_0 + \gamma'_z \mathbf{z}_{it})$, $\hat{\lambda}_{it}^{r(0)} = \lambda^r(\hat{\gamma}_0 + \hat{\gamma}'_z \mathbf{z}_{it})$, for the sampled farms using CMP r , for $r \in \{0, 1\}$.

(B.1) Compute the LS estimates of $(\alpha_{k,0}^r, \alpha_{k,w}^r, \omega_{k,ue}^r)$, $(\hat{\alpha}_{k,0}^r, \hat{\alpha}_{k,w}^r, \hat{\omega}_{k,ue}^r)$, by regressing $\ln x_{k,it}$ on $(1, \mathbf{w}_{it}, \hat{\lambda}_{it}^{r(0)})$ based on the sup-sample of farms using CMP r , for $k \in K$ and $r \in \{0, 1\}$.

(B.2) Compute the estimates of error terms $u_{k,it}^r$, $\hat{u}_{k,it}^r = \ln x_{k,it} - \hat{\alpha}_{k,0}^r - \mathbf{w}'_{it} \hat{\alpha}_{k,w}^r$, for the sampled farms using CMP r , for $k \in K$ and $r \in \{0, 1\}$. Construct the estimate of

¹⁸Calculation details leading to this result can be found in Appendix 1.6.1.

vector ω_{ue}^r , $\hat{\omega}_{ue}^r = (\hat{\omega}_{k,ue}^r : k \in K)$, for $r \in \{0, 1\}$.

(C.1) Compute the estimates of parameters $\omega_{k\ell,uu}^r$, with

$$\hat{\omega}_{k\ell,uu}^r = \left(\sum_{i,t} 1(r_{it} = r) \right)^{-1} \left(\sum_{i,t} 1(r_{it} = r) \hat{u}_{k,it}^r \hat{u}_{\ell,it}^r + \hat{\omega}_{k,ue}^r \hat{\omega}_{\ell,ue}^r \sum_{i,t} (\hat{\gamma}_0 + \hat{\gamma}'_z \mathbf{z}_{it}) \hat{\lambda}_{it}^{r(0)} (\gamma_0 + \gamma'_z \mathbf{z}_{it}) \right)$$

for $(k, \ell) \in K \times K$ and $r \in \{0, 1\}$.

(C.2) Construct the estimate of matrix Ω_{uu}^r , $\hat{\Omega}_{uu}^r = [\hat{\omega}_{k\ell,uu}^r : (k, \ell) \in K \times K]$, that of term ψ_{ee}^r , $\hat{\psi}_{ee}^r = 1 - (\hat{\omega}_{ue}^r)' (\hat{\Omega}_{uu}^r)^{-1} \hat{\omega}_{ue}^r$, and that of vector ρ_{eu}^r , $\hat{\rho}_{eu}^r = (\hat{\Omega}_{uu}^r)^{-1} \hat{\omega}_{ue}^r$, for $r \in \{0, 1\}$.

(C.3) Compute the estimates of control terms λ_{it}^r ,

$$\lambda_{it}^r, \hat{\lambda}_{it}^r = (\hat{\psi}_{ee}^r)^{-1/2} \lambda^r \left((\hat{\psi}_{ee}^r)^{-1/2} [\hat{\gamma}_0 + \hat{\gamma}'_z \mathbf{z}_{it} + (\hat{\rho}_{eu}^r)' \hat{u}_{it}^r] \right), \text{ for the sampled farms using CMP } r, \text{ for } r \in \{0, 1\}.$$

(D) Compute the nonlinear LS estimates of $(\beta^r, \rho_{vu}^r, \psi_{ve}^r)$, $(\hat{\beta}^r, \hat{\rho}_{vu}^r, \hat{\psi}_{ve}^r)$, by considering the approximate ‘‘doubly augmented’’ model of $\ln y_{it}$,

$$\ln y_{it} = \ln f^r(\mathbf{x}_{it}, \mathbf{c}_{it}; \beta^r) + (\rho_{vu}^r)' \hat{u}_{it}^r + \psi_{ve}^r \hat{\lambda}_{it}^r + \hat{u}_{it}^r$$

with $E[\hat{u}_{it}^r] = 0$, based on the sub-sample of farms using CMP r , for $r \in \{0, 1\}$.

(E) Use resampling techniques for computing the empirical distribution of the corresponding estimators of parameter vector $(\beta^r, \rho_{vu}^r, \psi_{ve}^r)$, as well as of parameter vector $(\alpha_{k,0}^r, \alpha_{k,w}^r, \omega_{k,ue}^r)$ for $k \in K$, for $r \in \{0, 1\}$.

1.4.3 Identifying assumptions, model nonlinearity and other estimation methods

The conditional joint normality assumptions given in Equation (1.10) are necessary to ensure the consistency of the results obtained with the above estimation procedure. Indeed, they are the foundations of the control functions we use to account for our multiple endogeneity issues. Relaxing the independence assumption of error term vectors $(v_{it}^r, \mathbf{u}_{it}^r, e_{it})$ and \mathbf{z}_{it} is very difficult, excepted for allowing heteroscedasticity of the error terms conditionally on \mathbf{z}_{it} . Relaxing the normality assumptions for vectors $(v_{it}^r, \mathbf{u}_{it}^r, e_{it})$ would require substantial adjustments in the estimation procedure presented above.

Another consistent estimation process for CMP specific yield models can be derived from Wooldridge (2010). The estimation procedure relies on (i) a set of conditional mean linearity conditions given by $E[v_{it}^r | \mathbf{z}_{it}, e_{it}] = \theta_v^r e_{it}$ and $E[\mathbf{u}_{it}^r | \mathbf{z}_{it}, e_{it}] = \theta_u^r e_{it}$ for $r \in \{0, 1\}$ and (ii) the normality assumption for the CMP choice model *i.e.* $e_{it} | \mathbf{z}_{it} \sim \mathcal{N}(0, 1)$. Under these assumptions, augmented yield model with control function $\omega_{ve}^r \lambda^r (\gamma_0 + \gamma'_z \mathbf{z}_{it})$ can be estimated by 2SLS. Mills ratio terms $\lambda^r (\gamma_0 + \gamma'_z \mathbf{z}_{it})$ are included in the instrument set of the 2SLS estimator and permit to deal with the endogenous sample selection issues. Yet, because we consider nonlinear yield

models, estimators based on orthogonality conditions, such as nonlinear 2SLS, are more complex to estimate.¹⁹ GMM can be a solution to get better estimation results. Indeed, GMM relies on estimated instruments that are designed for making better use of the information content of instrumental variables than standard nonlinear 2SLS estimators.²⁰ Let's consider the yield functions from the second step of the approach proposed by Wooldridge (2010):

$$\ln y_{it} = \ln f^r(\mathbf{x}_{it}, \mathbf{c}_{it}; \boldsymbol{\beta}^r) + \omega_{ve}^r \lambda^r (\gamma_0 + \boldsymbol{\gamma}'_z \mathbf{z}_{it}) + \mu_{it}^r,$$

where $E[\mu_{it}^r | \mathbf{z}_{it}, r_{it} = r] = 0$. The efficient instrument corresponding to this model is given by:

$$\zeta^r(\mathbf{z}_{it}) = E[(\mu_{it}^r)^2 | \mathbf{z}_{it}, r_{it} = r]^{-1} \frac{\partial}{\partial(\boldsymbol{\beta}^r, \omega_{ve}^r)} E[\ln f^r(\mathbf{x}_{it}, \mathbf{c}_{it}; \boldsymbol{\beta}^r) + \omega_{ve}^r \lambda^r (\gamma_0 + \boldsymbol{\gamma}'_z \mathbf{z}_{it}) | \mathbf{z}_{it}, r_{it} = r],$$

or, equivalently by:

$$\zeta^r(\mathbf{z}_{it}) = V[\mu_{it}^r | \mathbf{z}_{it}, r_{it} = r]^{-1} \begin{bmatrix} \mathbf{f}^r(\mathbf{z}_{it}; \boldsymbol{\beta}^r) \\ \lambda^r (\gamma_0 + \boldsymbol{\gamma}'_z \mathbf{z}_{it}) \end{bmatrix},$$

where $\mathbf{f}^r(\mathbf{z}_{it}; \boldsymbol{\beta}^r) = E \left[\frac{\partial}{\partial \boldsymbol{\beta}^r} \ln f^r(\mathbf{x}_{it}, \mathbf{c}_{it}; \boldsymbol{\beta}^r) | \mathbf{z}_{it}, r_{it} = r \right]$. In what follows, the conditional heteroskedasticity correction term $V[\mu_{it}^r | \mathbf{z}_{it}, r_{it} = r]^{-1}$ is ignored, as it is usually the case in practice. When $\ln f^r(\cdot)$ is linear in $\ln \mathbf{x}_{it}$ (e.g., standard Cobb-Douglas or Translog production functions), we have for the following formula for the gradient term:

$$\mathbf{f}^r(\mathbf{z}_{it}; \boldsymbol{\beta}^r) = (1, \mathbf{c}_{it}, E[\ln \mathbf{x}_{it} | \mathbf{z}_{it}, r_{it} = r]).$$

In this case, (near)-efficient instrument $\zeta^r(\mathbf{z}_{it})$ can easily be estimated *a priori*. It suffices to observe that the considered ERS model yields $E[\ln x_{k,it} | \mathbf{z}_{it}, r_{it} = r] = \alpha_{k,0}^r + \mathbf{w}'_{it} \boldsymbol{\alpha}_{k,w}^r + \omega_{k,ue}^r \lambda^r (\gamma_0 + \boldsymbol{\gamma}'_z \mathbf{z}_{it})$. Yet, in cases where $\ln f^r(\cdot)$ is nonlinear in $\ln \mathbf{x}_{it}$, computing gradient term $\mathbf{f}^r(\mathbf{z}_{it}; \boldsymbol{\beta}^r)$ is much more challenging. Given that $E[\ln \mathbf{x}_{it} | \mathbf{z}_{it}, r_{it} = r]$ can be computed, a possible solution consists of giving a rough approximation for the gradient term as

$$\mathbf{f}^r(\mathbf{z}_{it}; \boldsymbol{\beta}^r) \approx \frac{\partial}{\partial \boldsymbol{\beta}^r} \ln f^r (E[\ln \mathbf{x}_{it} | \mathbf{z}_{it}, r_{it} = r], \mathbf{c}_{it}; \boldsymbol{\beta}^r),$$

for determining instruments for estimating the model of y_{it} in the GMM framework. The structure of efficient instruments depending heavily on the functional form of the model, we cannot give a general result here.

¹⁹For instance, Latruffe et al. (2017) report estimation results that document this point. Standard nonlinear 2SLS estimators perform poorly when estimating their nonlinear stochastic production frontier models.

²⁰See, e.g., Chamberlain (1987) and Newey (1990, 1993). The considered instruments need to be sufficiently close to the efficient instrument of the considered estimation problem, the form of which was determined by Chamberlain (1987). They can be built based on preliminary estimation steps. Latruffe et al. (2017) report that GMM estimators based on suitably designed instruments substantially outperform standard 2SLS estimators when estimating nonlinear stochastic production frontiers.

On the contrary, easy implementation is one of the benefit of our approach. Indeed, our approach only involves using nonlinear LS and Probit ML estimators. These estimators are efficiently coded in standard statistical software and perform generally well in practice. Another advantage of our approach is its parsimony. A single control function suffices to deal with the correlation of endogenous explanatory variable with an error term, regardless of how this variable enters in the considered model (*e.g.*, linearly, polynomial, in logged form, *etc.*). The cost of such a simple estimation approach is that it relies on joint normality assumptions that are stronger than those considered by Wooldridge (2010).

1.5 Discussion

In this paper, we presented a novel estimation approach for an “extended” endogenous regime switching model with endogenous regressors. Previous approaches were unsatisfactory for our purposes. Because the model we considered entails nonlinearities, standard instrumental variable approaches are difficult to implement. We thus consider an approach relying on control functions to control for both endogenous regressors and regime. We develop an estimation approach inspired by Heckman two-step approach for selection model with a corrected inverse Mills ratio. The great asset of such estimation approach is its simple implementation combined with its parcimony. Yet, it comes at the price of *(i)* a lower efficiency compared to the ML approach and *(ii)* more restrictive distributional assumptions compared to the GMM approach. An interesting extension of this work would be to evaluate how much findings differ when using each estimation approach.

An empirical application of the extended ERS model to the case of endogenous covariates is provided in next chapter.

1.6 Appendices

1.6.1 Calculation details

Calculation details of Equation (1.13)

This appendix demonstrates the following expression of conditional expectation $E[\ln y_{it} | \mathbf{x}_{it}, \mathbf{z}_{it}, \mathbf{u}_{it}^r, r_{it} = r]$:

$$E[\ln y_{it} | \mathbf{x}_{it}, \mathbf{z}_{it}, \mathbf{u}_{it}^r, r_{it} = r] = \ln f^r(\mathbf{x}_{it}, \mathbf{c}_{it}; \boldsymbol{\beta}^r) + (\boldsymbol{\rho}_{vu}^r)' \mathbf{u}_{it}^r \\ + \psi_{ve}^r (\psi_{ee}^r)^{-1/2} \lambda^r \left((\psi_{ee}^r)^{-1/2} [\gamma_0 + \boldsymbol{\gamma}'_z \mathbf{z}_{it} + (\boldsymbol{\rho}_{eu}^r)' \mathbf{u}_{it}^r] \right).$$

It is easily shown that:

$$\begin{cases} E[v_{it}^r | \mathbf{x}_{it}, \mathbf{z}_{it}, \mathbf{u}_{it}^r, r_{it} = r] &= E[v_{it}^r | \mathbf{z}_{it}, \mathbf{u}_{it}^r, r_{it} = r] \\ E[v_{it}^r | \mathbf{z}_{it}, \mathbf{u}_{it}^r, r_{it} = r] &= (\boldsymbol{\rho}_{vu}^r)' \mathbf{u}_{it}^r + \psi_{ve}^r \lambda_{it}^r \end{cases}$$

The joint normality of vectors $(\mathbf{u}_{it}^r, v_{it}^r, e_{it}) | \mathbf{z}_{it} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Omega}^r)$ stated in Equation (1.10) yields $(v_{it}^r, e_{it}) | (\mathbf{z}_{it}, \mathbf{u}_{it}^r) \sim \mathcal{N}(\mathbf{R}^r \mathbf{u}_{it}^r, \boldsymbol{\Psi}^r)$ where:

$$\mathbf{R}^r = \begin{bmatrix} (\boldsymbol{\rho}_{uv}^r)' \\ (\boldsymbol{\rho}_{ue}^r)' \end{bmatrix} = \begin{bmatrix} (\boldsymbol{\omega}_{uv}^r)' (\boldsymbol{\Omega}_{uu}^r)^{-1} \\ (\boldsymbol{\omega}_{ue}^r)' (\boldsymbol{\Omega}_{uu}^r)^{-1} \end{bmatrix}$$

and

$$\boldsymbol{\Psi}^r = \begin{bmatrix} \psi_{vv}^r & \psi_{ve}^r \\ \psi_{ve}^r & \psi_{ee}^r \end{bmatrix} = \begin{bmatrix} \omega_{vv}^r - (\boldsymbol{\omega}_{uv}^r)' (\boldsymbol{\Omega}_{uu}^r)^{-1} \boldsymbol{\omega}_{uv}^r & \omega_{ve}^r - (\boldsymbol{\omega}_{uv}^r)' (\boldsymbol{\Omega}_{uu}^r)^{-1} \boldsymbol{\omega}_{ue}^r \\ \omega_{ve}^r - (\boldsymbol{\omega}_{uv}^r)' (\boldsymbol{\Omega}_{uu}^r)^{-1} \boldsymbol{\omega}_{ue}^r & 1 - (\boldsymbol{\omega}_{ue}^r)' (\boldsymbol{\Omega}_{uu}^r)^{-1} \boldsymbol{\omega}_{ue}^r \end{bmatrix}.$$

These results imply that terms v_{it}^r and e_{it} can be decomposed as the following:

$$v_{it}^r = (\boldsymbol{\rho}_{uv}^r)' \mathbf{u}_{it}^r + \varepsilon_{v,it}^r \\ e_{it} = (\boldsymbol{\rho}_{ue}^r)' \mathbf{u}_{it}^r + \varepsilon_{e,it}^r$$

where $(\varepsilon_{v,it}^r, \varepsilon_{e,it}^r) | (\mathbf{z}_{it}, \mathbf{u}_{it}^r) \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Psi}^r)$.

This also implies that:

$$E[v_{it}^r | \mathbf{z}_{it}, \mathbf{u}_{it}^r, r_{it} = r] = (\boldsymbol{\rho}_{vu}^r)' \mathbf{u}_{it}^r + E[\varepsilon_{v,it}^r | \mathbf{z}_{it}, \mathbf{u}_{it}^r, r_{it} = r].$$

Observing that $r_{it} = 1[\gamma_0 + \boldsymbol{\gamma}'_z \mathbf{z}_{it} + (\boldsymbol{\rho}_{ue}^r)' \mathbf{u}_{it}^r + \varepsilon_{e,it}^r \geq 0]$, it suffices to apply standard results on the means of truncated normal variables for obtaining

$$E[\varepsilon_{v,it}^r | \mathbf{z}_{it}, \mathbf{u}_{it}^r, r_{it} = r] = \psi_{ve}^r (\psi_{ee}^r)^{-1/2} \lambda^r \left((\psi_{ee}^r)^{-1/2} (\gamma_0 + \boldsymbol{\gamma}'_z \mathbf{z}_{it} + (\boldsymbol{\rho}_{eu}^r)' \mathbf{u}_{it}^r) \right) = \psi_{ve}^r \lambda_{it}^r.$$

Calculation details of Equation (1.17)

This appendix demonstrates that covariance parameter $\omega_{k\ell,uu}^r$ can be expressed as:

$$\omega_{k\ell,uu}^r = E[u_{k,it}^r u_{\ell,it}^r | r_{it} = r] + \omega_{k,ue}^r \omega_{\ell,ue}^r E[(\gamma_0 + \boldsymbol{\gamma}'_z \mathbf{z}_{it}) \lambda^r (\gamma_0 + \boldsymbol{\gamma}'_z \mathbf{z}_{it}) | r_{it} = r].$$

Given that $e_{it} | \mathbf{z}_{it} \sim \mathcal{N}(0, 1)$ and $r_{it} = 1[\gamma_0 + \boldsymbol{\gamma}'_z \mathbf{z}_{it} + e_{it} \geq 0]$, it is well known that:

$$E[e_{it} | \mathbf{z}_{it}, r_{it} = r] = \lambda^r (\gamma_0 + \boldsymbol{\gamma}'_z \mathbf{z}_{it}),$$

and

$$V[e_{it} | \mathbf{z}_{it}, r_{it} = r] = 1 - (\gamma_0 + \boldsymbol{\gamma}'_z \mathbf{z}_{it}) \lambda^r (\gamma_0 + \boldsymbol{\gamma}'_z \mathbf{z}_{it}) - \lambda^{2r} (\gamma_0 + \boldsymbol{\gamma}'_z \mathbf{z}_{it})^2,$$

which yields:

$$E[(e_{it})^2 | \mathbf{z}_{it}, r_{it} = r] = 1 - (\gamma_0 + \boldsymbol{\gamma}'_z \mathbf{z}_{it}) \lambda^r (\gamma_0 + \boldsymbol{\gamma}'_z \mathbf{z}_{it}).$$

The joint normality of vectors $(\mathbf{u}_{it}^r, v_{it}^r, e_{it}) | \mathbf{z}_{it} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Omega}^r)$ stated in Equation (1.10) allows to write \mathbf{u}_{it}^r as $\mathbf{u}_{it}^r = \boldsymbol{\omega}_{ue}^r e_{it} + \boldsymbol{\varepsilon}_{u,it}^r$, where $\boldsymbol{\varepsilon}_{u,it}^r | (\mathbf{z}_{it}, e_{it}) \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Omega}_{uu}^r - \boldsymbol{\omega}_{ue}^r (\boldsymbol{\omega}_{ue}^r)')$.

Note the last result shows that terms $\boldsymbol{\varepsilon}_{u,it}^r$ and e_{it} are independent conditionally on \mathbf{z}_{it} , implying in turn that residual terms $\boldsymbol{\varepsilon}_{u,it}^r = (\varepsilon_{k,u,it}^r : k \in K)$ do not depend on r_{it} conditionally on \mathbf{z}_{it} . These results also yield that

$$\begin{cases} u_{k,it}^r u_{\ell,it}^r = \omega_{k,ue}^r \omega_{\ell,ue}^r (e_{it})^2 + \omega_{k,ue}^r e_{it} \varepsilon_{\ell,u,it}^r + \omega_{\ell,ue}^r e_{it} \varepsilon_{k,u,it}^r + \varepsilon_{k,u,it}^r \varepsilon_{\ell,u,it}^r \\ E[e_{it} \varepsilon_{\ell,u,it}^r | \mathbf{z}_{it}, r_{it} = r] = E[\varepsilon_{\ell,u,it}^r] E[e_{it} | \mathbf{z}_{it}, r_{it} = r] = 0 \\ E[\varepsilon_{k,u,it}^r \varepsilon_{\ell,u,it}^r | \mathbf{z}_{it}, r_{it} = r] = E[\varepsilon_{k,u,it}^r \varepsilon_{\ell,u,it}^r] = \omega_{k\ell,uu}^r - \omega_{k,ue}^r \omega_{\ell,ue}^r \end{cases}.$$

Collecting these results gives:

$$E[u_{k,it}^r u_{\ell,it}^r | \mathbf{z}_{it}, r_{it} = r] = \omega_{k,ue}^r \omega_{\ell,ue}^r E[(e_{it})^2 | \mathbf{z}_{it}, r_{it} = r] + \omega_{k\ell,uu}^r - \omega_{k,ue}^r \omega_{\ell,ue}^r,$$

and, finally

$$E[u_{k,it}^r u_{\ell,it}^r | \mathbf{z}_{it}, r_{it} = r] = \omega_{k\ell,uu}^r - \omega_{k,ue}^r \omega_{\ell,ue}^r (\gamma_0 + \boldsymbol{\gamma}'_z \mathbf{z}_{it}) \lambda^r (\gamma_0 + \boldsymbol{\gamma}'_z \mathbf{z}_{it}).$$

Chapter 2

Estimation of production functions in low- and high-input production practices¹

2.1 Introduction

The need for a more sustainable agriculture is pressing and the reduction of adverse environmental and health effects of agriculture is on top of the agenda of food-value chain actors in Europe (Möhring et al., 2020). While current agricultural practices often have deleterious effects on the environment and on biodiversity, there is a need to maintain high production levels (see Foley et al., 2005, 2011). Low-input production practices have been proposed as a compromise between environmental and population growth concerns as they are scalable and thus applicable for large shares of production practices (Aune, 2012; Fess, Kotcon, and Benedito, 2011; Meynard, 2008). Low-input production practices are characterized by the use of adjusted management, *e.g.*, using agronomic principles and new technologies (*e.g.*, new varieties) that substantially reduce the use of synthetic pesticides and/or mineral fertilizers (see, *e.g.*, Bertrand and Doré, 2008; Meynard and Girardin, 1991; Rolland et al., 2003). Yet, the switch to low-input production practices comes with challenges for farmers and policy makers. First, the switch in production practices might be costly and result in more uncertain payoffs, in particular because yields are more volatile (*e.g.*, Finger, 2014; Gardebroek, Chavez, and Lansink, 2010; Greiner, Patterson, and Miller, 2009; Matyjaszczyk, 2019). Second, establishing political support may suffer from windfall effects, possibly reducing the efficacy of such interventions (Finger and El Benni, 2013; Pedersen et al., 2012). Public policy simulations accounting for farmers' practices can help to design policies that are more efficient. However, the effects of inputs on production and production risk might significantly differ across production systems. This may render some policies (*e.g.*, taxes) less efficient for some groups of farmers as marginal abatement costs might significantly differ across production systems. It is therefore important to map these differences and find suitable methods for their identification.

¹This chapter is written in collaboration with Niklas MÖHRING and Robert FINGER.

Earlier research on low-input production systems, was especially focused on agronomic aspects, often targeting organic production techniques. For example, to assess the sustainability of low-input practices some studies investigated into the soil biology and its microbial composition (*e.g.*, Clark et al., 1998; Kong et al., 2011; Mäder et al., 2000; Scow et al., 1994) and others looked at the impact in terms of yields (*e.g.*, *e.g.*, Clark et al., 1999; Liebhardt et al., 1989; Seufert, Ramankutty, and Foley, 2012). Economic research in this field addresses the differences between the low-input and high-input agriculture in terms of production and production risks (*e.g.*, Gardebroek, 2006; Gardebroek, Chavez, and Lansink, 2010; Serra, Zilberman, and Gil, 2008), efficiency and performance (*e.g.*, Acs et al., 2007; Lansink, Pietola, and Bäckman, 2002) as well as trade-offs with respect to the environmental performance (*e.g.*, Seufert and Ramankutty, 2017; Seufert, Ramankutty, and Foley, 2012). If self-selection effects and selection biases are well-known when dealing with technology choice (*e.g.*, Abdulai and Huffman, 2014; Alene and Manyong, 2007; Asfaw et al., 2012), most studies considering the impact of technology on agricultural production do not investigate on the differentiated role of inputs in production function. Yet, not considering these differences may lead to large biases in estimates.

We aim to fill this gap by estimating and comparing the production functions of high-input and low-input production practices. We consider a case study of Swiss wheat production, where both a low- and high-input wheat production system exist in parallel. More specifically, our empirical analysis uses rich, high resolution panel data on Swiss wheat production (617 observations, from 2009 to 2015), containing detailed information on output and input use (*i.e.* on pesticide use, mechanical weed control, fertilizer use, work and machinery), obtained from field journals. We consider estimating a primal production function while accounting for self-selection effects and selection bias. To this end, we consider an endogenous regime switching (ERS) model for low- and high-input technologies. To circumvent the problem of traditional instrumental variable (IV) estimation techniques with the significant nonlinearities in the damage abating part of the production function, we use control functions. It allows us to directly estimate (*i*) input use equations and (*ii*) different production functions for different farming types accounting for technology heterogeneity. Furthermore, identification of the differences between low-input and high-input farming through different parameters in the production and input use functions then allow us to directly estimate the effect of policies, *e.g.* due to bans of specific pesticides or taxes on inputs (Böcker, Möhring, and Finger, 2019; Finger et al., 2017).

Unfortunately, our empirical analysis suffers from weak instruments in the input use equations, implying that our results need to be interpreted cautiously. Nonetheless, our results tend to argue in favor of the production practice choice inducing

selection biases when estimating yield and input use models. The input-use demand model displays a positive estimated selection bias among high-input farmers, *i.e.* their unobserved characteristics seem to boost up their pesticide uses. On the other hand, the yield level model displays a positive estimated selection bias for low-input farmers, *i.e.* their unobserved characteristics allow them to achieve greater yield levels. Apart from that, the results from the yield model estimates tend also to show differentiated input effect between low- and high-input farmers. Yet, the weakness of the instrumental variables for input use encourages us to be very cautious in interpreting the size of the estimated coefficients in our production function. In particular, it challenges the estimation of the yield response to input use among both sub-samples. It does not detract from the need to consider the selection issue in order to evaluate the impact of the Swiss *Extenso* policy on pesticide uses.

The remainder of this article is structured as follows. The next section provides an overview of the introduction and adoption of low-input practices. The following section is dedicated to the methodological framework we used to explore both the input endogeneity issue and the self-selection resulting from the production practices choice. Once defined the methodological framework, we present the empirical model and the estimation strategy we use in this paper. Then, before presenting the results from the production function estimates, we first present the data. Presentation of the results is twofold. First, we give some insights from the estimation of the technology choice and input use models. Then, we present the results from the estimation of the production practices specific production function. Finally, results are discussed so we can draw conclusions from this work.

2.2 Background

2.2.1 The economic and agronomic principles of low-input production system

Agricultural policies in past decades encouraged intensive production practices, *e.g.* due to price support (Guay, 2012). Since the 1990s, support has been increasingly decoupled and low-input and integrated production systems have been increasingly supported. Integrated production relies on the integration of natural resources and regulation mechanisms to substitute for chemical inputs, *e.g.* through soil preparation operations, beneficial management practices or crop rotations. Farmers adopting low-input practices can also benefit from innovation such as crop breeding (Fess, Kotcon, and Benedito, 2011; Möhring et al., 2020; Rolland et al., 2003). Instead of being at the core of crop protection, pesticides in integrated pest management approaches are one element among others and should only be used when there is no other option left (Lucas, 2007). Hence a substantial reduction of the use of synthetic pesticides and mineral fertilizers can be achieved in these production practices (see, *e.g.*, Bertrand

and Doré, 2008; Meynard and Girardin, 1991; Rolland et al., 2003). In Switzerland, a voluntary integrated production program was introduced in 1992 (Boller et al., 1998; El Benni and Finger, 2013). Farmer participation in the program is incentivized with a price mark-up of around 5 CHF/100 kg (reflecting a *ca.* 10% price mark-up) and federal, direct payments of 400 CHF per hectare. These payments are conditional on farmers not using any fungicides, insecticides, plant growth regulators or chemical-synthetic stimulators of natural resistance. This low-input wheat production currently represents more than 50% of the total Swiss wheat production (Finger and El Benni, 2013). In this low-input scheme, fungicide and growth regulator use is substituted by variety choice and adjustment in production practices. Herbicides are mainly substituted via mechanical weed control. Insecticides play only a minor role in Swiss wheat production. Note that the use of mineral fertilizer is not restricted in the low-input scheme. Yet, the marginal effects of input use on wheat yields may differ across systems.

The impact of low-input production practices on production and the environment is multifaceted. Low-input systems tend to be associated with higher levels of organic carbon, soluble phosphorus, exchangeable potassium and potentially mineralizable nitrogen as well as a higher soil pH (Clark et al., 1998; Scow et al., 1994) and also impact the soil microbial mass (Mäder et al., 2000; Scow et al., 1994). Such changes in soil composition and biology are critical for long-term fertility maintenance and allow for a reduction in chemical input use. The global effect on crop-yield is quite complex to assess (Savary et al., 2019). Some studies reveal no significant difference on yields (Lechenet et al., 2017; Mischler et al., 2009) or argue that such differences are mainly contextual (Seufert, Ramankutty, and Foley, 2012). On the contrary, other studies demonstrate a decrease in yields for low-input production practices (Clark et al., 1999; Liebhardt et al., 1989). Yet, even if associated to a yield decrease, low-input agriculture can show similar returns to high-input agriculture because of lower costs for inputs and price premia (Clark et al., 1999; Rolland et al., 2003). However, low-input practices might be riskier than high-input practices (Ridier et al., 2013), lowering incentives for uptake by farmers and creating inertia in the change of production practices.

2.2.2 How to integrate technology in farmers' production function?

Standard production functions assume an homogeneous technology among farmers, thereby assume that inputs play the same role in the farmers' production functions, other heterogeneity sources being accounted for. The presence of low- and high-input production practices, which both rely on different agronomic principles, encourages us to consider heterogeneous technologies. Yet, integrate technology choice in production function is not trivial. If farmers face heterogeneous opportunity costs when

considering technology adoption, arise then endogenous selection issues - in particular, a self-selection effect here - when comparing new and high-input technologies. On that matter, Finger and El Benni (2013) found that adopters of low-input practices tend to be “smaller” farmers with already “lower” yields, which may indicate that opportunity costs of adopting low-input practices are lower for smaller farms. Additionally, the choice to adopt practices with more uncertain pay-offs might also reveal differences in farmers’ attitudes. Mzoughi (2011) shows that farmers adopting integrated crop protection tend to value the protection of the environment. When they are not controlled for, such factors that affect both technology choice and other production choices and outcomes induce endogeneity issues that are known in this case as selection bias.² Accounting for such selection bias is fundamental as such unobserved factors might also affect farmers’ responses to public policy incentives (Finger and El Benni, 2013; Mzoughi, 2011).

Endogenous regime switching models are the most common strategy to account for the effects of technology in the agricultural production literature as they address these selection bias issues. This approach was adopted for instance by Alene and Manyong (2007) to evaluate the impact of farmer education on the productivity with traditional and improved technology. Endogenous regime switching models were also used by Asfaw et al. (2012) and Abdulai and Huffman (2014).³ Yet, unlike most studies, we aim at studying the effect of technology on the response of crop yield to chemical input uses, in particular to pesticides. This requires the use of primal production function and as a consequence handling the issue of input uses endogeneity. Thus, the need to extend the standard endogenous regime switching model for addressing both technology choice and input use endogeneity issues. Plus, in order to correctly represents the steep decrease in crop protection level induced by reducing pesticide uses, the use of a damage abatement function as proposed by Lichtenberg and Zilberman (1986) seems appropriate. Hence, we consider a production function that is separated between two parts: a potential yield function and a damage abatement function. The potential yield function describes how productive inputs contribute to the potential crop yield level, *i.e.* the yield level that is free of any damage due to weeds, pests and/or diseases. The damage abatement function gives the share of potential yield that is saved by using pesticides, which are chemical crop protection inputs. To account for the potential complementary of inputs between them, we allow inputs to have both a productive and damage abating role on crop

²Many common factors impact both technology choices on the one hand, and production choices and outcomes on the other hand. If many of these factors are observed (*e.g.* market prices, weather conditions or general farmer characteristics), many of them are unobserved – including soil quality, farmer skills or their environmental preferences.

³Other approaches such as propensity score matching (Mayen, Balagtas, and Alexander, 2010) eventually combined to a difference-in-differences (DiD) approach (Greene and Liu, 2021; Mennig and Sauer, 2020) can be noted. Yet, they are not adapted in our context, especially the DiD approach, as we do not have data on the yield levels of farmers before the implementation of the Swiss policy encouraging extensive practices.

yields (Saha, Shumway, and Havenner, 1997).

The next two sections are dedicated first to the presentation of the econometric framework and how it is implemented.

2.3 Methodological framework: from a standard ERS to an "extended" ERS model

We first introduce the standard endogenous regime switching (ERS) model framework that we use to estimate our production and input demand functions. Let y_{it}^r , respectively $\mathbf{x}_{it}^r = (x_{k,it}^r : k \in K)$, denote the wheat yield level, respectively the input use levels, of farmer i in year t under the condition that this farmer uses production practice $r \in \{0, 1\}$. Let \mathbf{c}_{it} denote control variables that might impact production choices (*e.g.*, topographic and weather conditions) and \mathbf{p}_{it} be the vector of relevant prices (*e.g.*, crop and input prices). We assume that yield and input use levels can be modelled the following system of yield and input demand functions that are specific to the considered production practices:

$$\begin{cases} \ln y_{it}^r &= \ln f^r(\mathbf{x}_{it}^r, \mathbf{c}_{it}; \boldsymbol{\beta}^r) + v_{it}^r \\ \ln x_{k,it}^r &= \alpha_{k,0}^r + \mathbf{w}'_{it} \boldsymbol{\alpha}_{k,w}^r + u_{k,it}^r \text{ for } k \in K \end{cases} \quad (2.1)$$

for $r \in \{0, 1\}$ and where $\mathbf{w}_{it} = (\mathbf{p}_{it}, \mathbf{c}_{it})$. Function $f^r(\cdot)$ represents the general form of the production function for production practice r . Vector $\boldsymbol{\beta}^r$, respectively $\boldsymbol{\alpha}_k^r = (\alpha_{k,0}^r, \boldsymbol{\alpha}_{k,w}^r)$, is the parameter vector of the yield, respectively of input k , model. Finally, term v_{it}^r , respectively $u_{k,it}^r$, is the error term of the yield model, respectively of input k demand model.

We define the production technology choice model as a standard linear index model:

$$r_{it} = \mathbb{1}(\gamma_0 + \boldsymbol{\gamma}'_z + e_{it} \geq 0). \quad (2.2)$$

In vector $\mathbf{z}_{it} = (\mathbf{w}_{it}, \mathbf{q}_{it})$, vector \mathbf{q}_{it} collects farm(er) characteristics that might affect production practices choice without directly impacting production (*e.g.*, age and education). Vector \mathbf{z}_{it} is used as a vector of instrumental variables for endogenous technology choice r_{it} . This dichotomous choice model depends on parameter vector $\boldsymbol{\gamma} = (\gamma_0, \boldsymbol{\gamma}_z)$ and contains an error term, e_{it} . Condition $r_{it} = 1$ states that farmer i chooses the low-input production technology at time t , while condition $r_{it} = 0$ states that s/he chooses the high-input production technology.

We assume that the variables collected in vector \mathbf{z}_{it} , which includes \mathbf{w}_{it} , are exogenous in the models given in Equations 2.1 and 2.2. But, we assume that input uses \mathbf{x}_{it}^r can be endogenous in the production function of y_{it}^r , implying that error terms v_{it}^r and $\mathbf{u}_{it}^r = (u_{k,it}^r : k \in K)$ may be correlated. Finally, e_{it} represents the error term of

the technology choice model. We also assume that the production practices choice can be endogenous in the production and input demand equations, implying that error terms e_{it} and $(v_{it}^r, \mathbf{u}_{it}^r)$ may be correlated. Indeed, production practices, input use and yield levels choices may all be impacted by common unobserved factors.

Since the seminal work of Heckman (1976, 1979) and Lee (1978), most ERS models considered in applied econometric studies are parametric models based on multivariate normality assumptions. Equations 2.1 and 2.2 define an ERS model that describes a vector of production choices (yield and input use levels), the models of which depend on a regime (technology) choice. We extend the usual distributional assumptions on the error term vectors by assuming that vectors $(v_{it}^r, \mathbf{u}_{it}^r, e_{it})$ are multivariate normal and independent of the model information set, \mathbf{z}_{it} :

$$(\mathbf{u}_{it}^r, v_{it}^r, e_{it}) | \mathbf{z}_{it} \sim N(\mathbf{0}, \mathbf{\Omega}^r) \text{ with } \mathbf{\Omega}^r = \begin{bmatrix} \mathbf{\Omega}_{uu}^r & \mathbf{\Omega}_{uv}^r & \mathbf{\Omega}_{ue}^r \\ (\mathbf{\Omega}_{uv}^r)' & \omega_{vv}^r & \omega_{ve}^r \\ (\mathbf{\Omega}_{ue}^r) & \omega_{ve}^r & 1 \end{bmatrix} \text{ for } r \in \{0, 1\} \quad (2.3)$$

where

$$\mathbf{\Omega}_{uv}^r = (\omega_{k,uv}^r : k \in K), \mathbf{\Omega}_{ue}^r = (\omega_{k,ue}^r : k \in K) \text{ and } \mathbf{\Omega}_{uu}^r = [\omega_{k\ell,uu}^r : (k, \ell) \in K \times K].$$

Our assumption stating that error term e_{it} is normal, with $e_{it} | \mathbf{z}_{it} \sim \mathcal{N}(0, 1)$, implies that the production practices choice model is a standard Probit model. The normalization restriction stating that $V[e_{it}] = 1$ is standard in this context.

Importantly, the Gaussian ERS model given in Equations (2.1) – (2.3) is not standard since input use levels may be endogenous in the corresponding technology specific production functions. Input use endogeneity in production function estimation problems is a long-standing issue that has received much attention in the econometric literature (*e.g.*, Akerberg, Caves, and Frazer, 2015). Yet, to our knowledge this issue has not been considered in modelling framework involving ERS models.

2.4 Econometric implementation

2.4.1 A Cobb-Douglas crop production function with a damage abatement part

First, let describe the functional form we use for our production function. Apart from using a primal function, we consider a crop production function where input uses are assumed to be asymmetric. Thus, we consider three type of inputs: (i) purely productive inputs that increase yields, (ii) purely damage abating inputs that decrease the impact of damage events on yields, and (iii) interactive inputs that might suffer from complementary with other inputs and thus cannot be classified as

purely productive (or purely damage abating). Let denote $h^r(\cdot)$ (respectively $g^r(\cdot)$) the productive part (respectively the damage abating part) of the production function and $\mathbf{x}_{(h)}$ (respectively $\mathbf{x}_{(g)}$) the productive and interactive inputs (respectively the damage abating and interactive inputs) that are covariates in function $h^r(\cdot)$ (respectively $g^r(\cdot)$). We can thus rewrite the production practices specific crop yield function as:

$$f^r(\mathbf{c}, \mathbf{x}; \boldsymbol{\beta}^r) = s^r(\mathbf{c}; \boldsymbol{\beta}_{(c)}^r) h^r(\mathbf{x}_{(h)}; \boldsymbol{\beta}_{(h)}^r) g^r(\mathbf{x}_{(g)}; \boldsymbol{\beta}_{(g)}^r), \quad (2.4)$$

where $s^r(\cdot)$ defines the impacts of production conditions and farms' characteristics \mathbf{c} on crop yields. $(\boldsymbol{\beta}_{(c)}^r, \boldsymbol{\beta}_{(h)}^r, \boldsymbol{\beta}_{(g)}^r)$ is the set of production practices specific coefficients from the production function.

Following Zhengfei et al. (2005, 2006) and Möhring et al. (2020), our empirical crop yield function models combine Cobb-Douglas potential yield functions

$$h(\mathbf{x}_{(h),it}; \boldsymbol{\beta}_{(h)}^r) = \exp(\ln \mathbf{x}'_{(h),it} \boldsymbol{\beta}_{(h),x}^r),$$

and quadratic damage abatement functions

$$g^r(\mathbf{x}_{(g),it}; \boldsymbol{\beta}_{(g)}^r) = \exp\left(-(\beta_{(g),0}^r + \mathbf{x}'_{(g),it} \boldsymbol{\beta}_{(g),x}^r)^2\right).$$

Quadratic damage abatement functions are convenient as negative estimates of $\boldsymbol{\beta}_{(g)}^r$ are allowed and thus no particular estimation restrictions are needed. As for the $s^r(\cdot)$ part of the production function, we use a simple linear form. From that we can derive the detailed yield production function as

$$\ln y_{it} = \beta_0^r + \mathbf{c}'_{it} \boldsymbol{\beta}_{(c)}^r + \ln \mathbf{x}'_{(h),it} \boldsymbol{\beta}_{(h),x}^r - (\beta_{(g),0}^r + \mathbf{x}'_{(g),it} \boldsymbol{\beta}_{(g),x}^r)^2 + v_{it}^r, \quad (2.5)$$

where v_{it}^r still represents the error term.⁴

2.4.2 An original, multistep, estimation procedure

Our ERS model being fully parametric, it could be estimated in the maximum likelihood (ML) framework. Yet, its involving yield functions with nonlinear functional form and its large parameter vector make the practical computation of maximum likelihood estimators challenging. The distributional assumptions given in Equation (2.3) enables us to define an alternative, multistage, estimation approach for our endogenous regime switching model. This estimation approach is fairly simple since it is defined as a sequence of estimation problems that are easy to solve, that is to say Probit model estimation and least squares (LS) problems.

⁴The log transformation is standard in the agricultural production literature. It is particularly convenient when using Cobb-Douglas functions as it permits to linearize the model. Yet, in our model, the log transformation does not permit to linearize the model due to the damage abatement function.

This estimation approach can be seen as an extension of Heckman's two-step approach for estimating standard (*i.e.* with exogenous explanatory variables) Gaussian ERS models (*e.g.*, Heckman, Tobias, and Vytlačil, 2003; Wooldridge, 2002, 2015) to the case with endogenous regressors. It relies on two sets of control functions. The first one is used to deal with the input use endogeneity issue in the production function while the second one is used to deal with the sample selection issues due to the production practices choice.

Let consider first the estimation of the input demand models. Standard results due to Heckman (1976, 1979) yield that:

$$E[u_{k,it}^r | \mathbf{z}_{it}, r_{it} = r] = \omega_{k,ue}^r \lambda^r(\gamma_0 + \boldsymbol{\gamma}'_z \mathbf{z}_{it}) \text{ for } k \in K \text{ and } r \in \{0, 1\}. \quad (2.6)$$

Function $\lambda^r(\cdot)$ is defined by

$$\lambda^r(a) = (2r - 1)\lambda((2r - 1)a) \text{ for } r \in \{0, 1\}.$$

Following standard notations, function $\lambda(s) = \Phi(s)^{-1}\varphi(s)$ denotes the inverse Mills ratio at a given that function $\varphi(s)$, respectively $\Phi(s)$, denotes the probability, respectively cumulative, distribution function of $\mathcal{N}(0, 1)$ at s . Equation (2.6) is based on the joint normality of error term pairs $(u_{k,it}^r, e_{it})$ implied by assumption (2.3), on technology choice model (2.2) and on the input use models given in Equation (2.1).

Equation (2.6) yields that equation:

$$\ln x_{k,it} = \alpha_{k,0}^r + \mathbf{w}'_{it} \boldsymbol{\alpha}_{k,w}^r + \omega_{k,ue}^r \lambda^r(\gamma_0 + \boldsymbol{\gamma}'_z \mathbf{z}_{it}) + \eta_{k,it}^r, \quad (2.7)$$

where $E[\eta_{k,it}^r | \mathbf{z}_{it}, r_{it} = r] = 0$ defines regression models in the sub-samples defined by the production practices choices. Term $\omega_{k,ue}^r \lambda^r(\gamma_0 + \boldsymbol{\gamma}'_z \mathbf{z}_{it})$ defines a control function for endogenous selection of the observations characterized by $r_{it} = r$ in the considered sub-sample. Error term $\eta_{k,it}^r$ is defined by $\eta_{k,it}^r = u_{k,it}^r - \omega_{k,ue}^r \lambda^r(\gamma_0 + \boldsymbol{\gamma}'_z \mathbf{z}_{it})$, which implies that exogeneity condition $E[\eta_{k,it}^r | \mathbf{z}_{it}, r_{it} = r] = 0$ necessary holds (in the sub-sample characterized by $r_{it} = r$).⁵

Input demand functions can thus be estimated following Heckman's two-step approach. First, the Probit model of $r_{it} | \mathbf{z}_{it}$ (defined by Equation (2.2)) is estimated by ML for obtaining a consistent estimate of parameter $\boldsymbol{\gamma}$, $\hat{\boldsymbol{\gamma}} = (\hat{\boldsymbol{\gamma}}_0, \hat{\boldsymbol{\gamma}}_z)$. This estimate

⁵In practice, because of *zeros* among the pesticide input uses, we consider an inverse hyperbolic sine (IHS) transformation for pesticide use levels. We keep the log transformation for the nitrogen variable. As the IHS transformation of pesticide use only intervene in the input use demand models and as residuals in the yield control function, we expect a low impact on the estimation results of the yield production function. We consider the IHS transformation rather than deleting the zeros because among the 27 zeros observations, 26 belong to the high-input category. It means we would loose 10% of our high-input sub-sample, which is already smaller than the low-input one. Yet, a sensitivity analysis when deleting these *zeros* from the observations will be performed.

can then be used for obtaining consistent estimates of Mills ratio terms, $\lambda^r(\hat{\gamma}_0 + \hat{\gamma}'_z \mathbf{z}_{it})$. Second, consistent estimates of parameters $(\alpha^r_{k,0}, \alpha^r_{k,w}, \omega^r_{k,ue})$ can be obtained by applying standard linear LS to Equation (2.7) in the sub-sample characterized by $r_{it} = r$ after replacing terms $\lambda^r(\gamma_0 + \gamma'_z \mathbf{z}_{it})$ by their estimates. Let vectors $(\hat{\alpha}^r_{k,0}, \hat{\alpha}^r_{k,w}, \hat{\omega}^r_{k,ue})$ denote these parameter estimates.

Let now consider the estimation problem of the yield functions given in Equation (2.1). The functional form of the models of $\ln y_{it}$, assumption (2.3) and technology choice model (2.2) yield that:

$$E[\ln y_{it} | \mathbf{x}_{it}, \mathbf{z}_{it}, \mathbf{u}^r_{it}, r_{it} = r] = \ln f^r(\mathbf{x}_{it}, \mathbf{c}_{it}; \boldsymbol{\beta}^r) + (\boldsymbol{\rho}^r_{vu})' \mathbf{u}^r_{it} + \psi^r_{ve} \lambda^r_{it} \quad (2.8)$$

where

$$\lambda^r_{it} = (\psi^r_{ee})^{-1/2} \lambda^r \left((\psi^r_{ee})^{-1/2} (\gamma_0 + \gamma'_z \mathbf{z}_{it} + (\boldsymbol{\rho}^r_{eu})' \mathbf{u}^r_{it}) \right) \quad (2.9)$$

and

$$\begin{aligned} \boldsymbol{\rho}^r_{vu} &= (\boldsymbol{\Omega}^r_{uu})^{-1} \boldsymbol{\omega}^r_{uv}, \\ \boldsymbol{\rho}^r_{eu} &= (\boldsymbol{\Omega}^r_{uu})^{-1} \boldsymbol{\omega}^r_{ue}, \\ \psi^r_{ve} &= \omega^r_{ve} - (\boldsymbol{\omega}^r_{uv})' (\boldsymbol{\Omega}^r_{uu})^{-1} \boldsymbol{\omega}^r_{ue}, \\ \psi^r_{ee} &= 1 - (\boldsymbol{\omega}^r_{ue})' (\boldsymbol{\Omega}^r_{uu})^{-1} \boldsymbol{\omega}^r_{ue}. \end{aligned} \quad (2.10)$$

Equations (2.8)–(2.10), which are demonstrated in Chapter 1, describe the first key result underlying our estimation approach. In Equation (2.8), term $(\boldsymbol{\rho}^r_{vu})' \mathbf{u}^r_{it}$ is a control function for the endogeneity of input choices \mathbf{x}_{it} (with $\mathbf{x}_{it} = \mathbf{x}^r_{it}$) while term $\psi^r_{ve} \lambda^r_{it}$ is a control function for the selection of observations characterized by $r_{it} = r$. Error term \mathbf{u}^r_{it} can be defined as function of exogenous variables that include price vector \mathbf{p}_{it} , with $\mathbf{u}^r_{it} = \log x^r_{k,it} - \alpha^r_{k,0} - \mathbf{w}'_{it} \boldsymbol{\alpha}^r_{k,w}$. This vector plays the role of instrumental variables of the endogenous input use levels \mathbf{x}^r_{it} in the model of $\ln y^r_{it}$. Equation (2.9) shows that adding the usual Mills ratios, $\lambda^r(\gamma_0 + \gamma'_z \mathbf{z}_{it})$, is inappropriate for controlling for sample selection in the yield model with endogenous input uses. Term λ^r_{it} has error term \mathbf{u}^r_{it} as an argument and needs to be scaled by parameter $(\psi^r_{ee})^{-1/2}$. Equations (2.8) and (2.9) imply that a model of $\ln y_{it}$ can be written as

$$\ln y_{it} = \ln f^r(\mathbf{x}_{it}, \mathbf{c}_{it}; \boldsymbol{\beta}^r) + (\boldsymbol{\rho}^r_{vu})' \mathbf{u}^r_{it} + \psi^r_{ve} \lambda^r_{it} + \mu^r_{it}, \quad (2.11)$$

with $E[\mu^r_{it} | \mathbf{x}_{it}, \mathbf{z}_{it}, \mathbf{u}^r_{it}, \lambda^r_{it}, r_{it} = r] = 0$ and where μ^r_{it} is the residual term defined by $\mu^r_{it} = v^r_{it} - (\boldsymbol{\rho}^r_{vu})' \mathbf{u}^r_{it} - \psi^r_{ve} \lambda^r_{it}$. Equation (2.11) states that $\ln y_{it}$ follows a regression model in $(\mathbf{x}_{it}, \mathbf{c}_{it}, \mathbf{u}^r_{it}, \lambda^r_{it})$ conditionally on $r_{it} = r$. Accordingly, consistent estimates of parameters $\boldsymbol{\beta}^r$, $\boldsymbol{\rho}^r_{vu}$ and ψ^r_{ve} can be obtained by applying non linear LS to the “augmented” regression model given in Equation (2.11) based on the sub-sample of farmer using production practices r . Yet, implementing this approach supposes that error terms \mathbf{u}^r_{it} and Mills ratio terms λ^r_{it} can be estimated *a priori*. Consistent estimates of error terms vectors \mathbf{u}^r_{it} , denoted by $\hat{\mathbf{u}}^r_{it} = (\hat{u}^r_{k,it} : k \in K)$, are easily obtained by using the residual terms of the input demand models, that is to say terms

$$u_{k,it}^r = \ln x_{k,it}^r - \alpha_{k,0}^r + \mathbf{w}'_{it} \boldsymbol{\alpha}_{k,w}^r \text{ for } k \in K \text{ and } r \in \{0, 1\}.$$

The second key result on which our estimation is based implies that parameters ψ_{ee}^r and ρ_{eu}^r can be estimated *a priori*, implying that terms λ_{it}^r can also be estimated *a priori*. Indeed, given that we can consistently estimate ω_{ue}^r *a priori*, conditions on ψ_{ee}^r and ρ_{eu}^r in Equation (2.10) show that we just need to obtain consistent estimates of matrices $\boldsymbol{\Omega}_{uu}^r$ – i.e. parameters $\omega_{k\ell,uu}^r$ for $k \in K$ and $r \in \{0, 1\}$ – for obtaining consistent estimates of parameters ψ_{ee}^r and ρ_{eu}^r , and thus of terms λ_{it}^r . The following result, which is demonstrated in Chapter 1, provides moment conditions for estimating the elements of $\boldsymbol{\Omega}_{uu}^r$:

$$\omega_{k\ell,uu}^r = E[u_{k,it}^r u_{\ell,it}^r | r_{it} = r] + \omega_{k,ue}^r \omega_{\ell,ue}^r E[(\gamma_0 + \boldsymbol{\gamma}'_z \mathbf{z}_{it}) \lambda^r(\gamma_0 + \boldsymbol{\gamma}'_z \mathbf{z}_{it}) | r_{it} = r], \quad (2.12)$$

for $(k, \ell) \in K \times K$.

Conditional expectation $E[u_{k,it}^r u_{\ell,it}^r | r_{it} = r]$ is consistently estimated by its sample counterpart in the sub-sample characterized by $r_{it} = r$. All parameters involved in the right hand side term of Equation (2.12) can also be estimated *a priori*. As a result, the sample counterpart of the right hand side term of Equation (2.12) provides a consistent estimator of parameter $\omega_{k\ell,uu}^r$, implying that a consistent estimate of matrix $\boldsymbol{\Omega}_{uu}^r$ is easily obtained.

The results and insights given above yield an estimation procedure, which is described in Chapter 1, for estimating the ERS model with endogenous regressors given in Equations (2.1)–(2.3). The technology choice model is estimated first. Results of this estimation step provide elements for computing control functions that are used in the next step. The second estimation step consists of estimating the ERS input use models. Once again, results obtained in this estimation step are used for computing control functions to be used in the next step. The ERS yield function models are estimated in the last step. Each estimation step is fairly easy to implement in practice. Yet, computing the asymptotic distribution of the estimators obtained by using the considered multiple step estimation procedure is not straightforward. Bootstrap methods were used so to get the empirical standard errors of the obtained estimates (Efron and Tibshirani, 1986).

2.4.3 Robustness checks

Following Möhring et al. (2020), we consider work and machinery as purely productive inputs and fertilizers as interactive inputs. Effectively, fertilizers tend to increase weed and disease damages, thereby implying that pesticides and fertilizers can display complementary effects (Carpentier and Weaver, 1997). Yet, one can argue that work might also be an interactive input. In particular, tillage can impact potential yield by impacting the development of the crop root systems or impacting

soil composition. But tillage can also be used as a substitute for herbicides and thus can be considered as a damage abating input. To answer such concern, and more globally the concern expressed by Carpentier and Weaver (1997) on the relevance to use production function specifications considering separated roles of productive *versus* damage abating inputs, we perform asymmetry and separability tests (see Appendix 2.9.1).

Other robustness checks were performed. The first one considers global pesticide uses instead of detailed pesticide uses (*e.g.*, herbicides, fungicides, insecticides) in the yield function so to capture the global effect of pesticides on yield.⁶ Then, another robustness check considers the revenue variable as dependent variable instead of yield. Indeed, revenue variable as the advantage to integrate a quality component that is not captured by yield. We also consider a Translog function for the production function instead of the Cobb-Douglas. Indeed, the Translog function is often argued to be a more flexible functional form than Cobb-Douglas.

2.5 Data

We use an unbalanced panel dataset containing information on each management operation performed by the sampled farmers (*e.g.*, soil preparation, sowing, fertilization, crop protection operations and harvest) from 2009 to 2015. Data is obtained from Agroscope, the Swiss center of excellence for agricultural research and provided by the Swiss Central Evaluation of Agri-Environmental Indicators (Baan, Spycher, and Daniel, 2015). We transform initial data consisting of daily records on crop management and inputs use as follows (based on Möhring et al., 2020): we (*i*) calculate cost-equivalents for used machinery and working time, (*ii*) convert fertilizer applications into nitrogen equivalents and (*iii*) express pesticide use in terms of pesticide load. To this end, we use the Pesticide Load Index (Kudsk, Jørgensen, and Ørum, 2018; Möhring, Gaba, and Finger, 2019), which accounts for differences in standard dosages and the heterogeneous properties of pesticides.⁷ Möhring, Gaba, and Finger (2019) and Möhring et al. (2020) highlight the importance of pesticide indicator choice for the estimation of production effects. We combine this first dataset with farm-level bookkeeping data⁸ to obtain information on whether or not the farmer participate in the low-input production system for winter wheat (based on direct payments they receive).

⁶In fact, both estimation results are interesting because they tell a different story. Whereas the detailed pesticide specification permits to see the differentiated impact of each type of pesticides, the global pesticide specification allows to measure the overall effect of pesticides on yield as a damage abating input.

⁷See R Package *PesticideLoadIndicator*, available at <https://CRAN.R-project.org/package=PesticideLoadIndicator> for calculation details.

⁸This data is from retrieved Agroscope (Mouron and Schmid, 2012).

The data gathers 617 observations from 151 winter wheat farmers from 2009 to 2015.⁹ Table 2.1 provides the characteristics of high- and low-input farms in terms of input use and output. In addition to standard descriptive statistics, we performed Wilcoxon and Fligner-Killeen non-parametric tests for the equality of mean and variance, respectively.

TABLE 2.1: Descriptive statistics for low- and high-input Swiss wheat farmers

	Low-input Mean (<i>sd</i>)	High-input Mean (<i>sd</i>)	Mean test (p-value)	Var test (p-value)
Dependent variable (y_i)				
Yield (q/ha)	58.5 (8.9)	64.52 (10.26)	***	***
Revenue (CHF/ha)	3,361 (666)	2,869 (630)	***	0.16
Control variables (x_c)				
Mountain region	35%	31%	-	-
French speaking	30.8%	38%	-	-
Farm surface (ha)	26.16 (11.43)	30.78 (12.33)	***	***
Winter wheat surface	4.651 (3.615)	5.579 (3.760)	***	*
Winter wheat share	17.59% (10.07)	18.2% (10.72)	0.26	0.68
Weather variables (δ)				
Temperatures (°C)	9.251 (1.037)	9.191 (0.832)	0.14	***
log(Rainfall) (log(1000L m ⁻²))	6.965 (0.199)	6.974 (0.173)	0.30	**
Productive inputs (x)				
Work & machinery (CHF/ha)	1,430 (338)	1,697 (345.5)	***	*
Interactive inputs (s)				
Fertilizers (kg/ha)	141.8 (48.1)	147.2 (49.7)	***	0.71
Damage abating inputs (z)				
Mechanical pest control (CHF/ha)	314.6 (209)	291.4 (186)	0.26	**
Pesticides (LI/ha)	0.903 (1.998)	2.684 (2.027)	***	***
Herbicides (LI/ha)	0.859 (1.987)	0.871 (1.124)	***	***
Insecticides (LI/ha)	.	0.093 (0.640)	-	-
Fungicides (LI/ha)	.	1.575 (1.172)	-	-
	N = 381	N = 236		

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

LI = Load Index indicator

Source: Authors' calculations on Agroscope data.

As expected, farmers cultivating *low-input* wheat use fewer chemical inputs and face lower yields than farmers with intensive practices. Yet, their revenues per hectare are higher, because of the direct payment and the price premium they benefit from. When removing the direct payment (400 CHF/ha), low-input producers tend to have revenues similar to their high-input counterpart despite lower yield levels thanks to the price premium they benefit from. From control variables, we can see that characteristics of low- and high-input farmers tend to differ. Farms located in mountainous region are more frequent among low-input farms, confirming the findings of Finger and El Benni (2013) on free-rider effects of the program for farms with a lower yield

⁹Appendix 2.9.2 gives the frequency of observations from 2009-2015 for both sub-samples.

potential. Lastly, farms located in the French speaking part of Switzerland are over-represented among the high-input farmers. This may be due to differences in farm characteristics but also cultural differences and/or differences in extension services. This led us to incorporate a language dummy in our analyses. Weather variables were introduced as to control for the effects of meteorological event on yields and input uses. Weather conditions, average annual temperatures in °C and average annual rainfall in 1000 liter m⁽⁻²⁾ here are comparable between the two groups.

As said previously, price variables were used in the models for technology choice and input uses. Winter wheat prices can differ depending on the chosen variety. Low-input farmers benefit from a price premium of around 5 CHF/100 kg.¹⁰ This premium was not accounted for when using winter wheat prices. As for pesticide prices, the initial data we had were prices at the product level (price data from the largest input provider in Swiss agriculture, Fenaco, were used).¹¹ To control for the difference of recommended quantities between products, we computed pesticide costs per hectare. Then, we calculated the annual average of herbicide, insecticide and fungicide prices of all products that were considered in our data. This price proxy stands as long as we assume that the pesticide products used by the observed farmers are representative of what can be found on the pesticide market. Even if not representative of the overall pesticide prices, we can argue that the products used by our sample of farmers is representative of the winter wheat pesticide market. Yet, the fact that price was not available for all products might introduce a bias if data were not missing at random. When choosing technology and input uses, winter wheat prices (and input prices on a lesser extent) are not observed by farmers. Price anticipations are generally used in such situations and the most naïve anticipation is to use the observed prices at time $t - 1$. Yet, by doing so, we would have lost more than 10% of our sample as 2009 is the year when we have the greatest number of observations. Steadiness of prices over the period encouraged us to use current year price as price anticipation. As for fertilizer prices, we use national representative data from Agristat on costs of nitrogen fertilizers in CHF/ha (Agristat, 2009-2015).¹²

¹⁰The actual price mark-up paid to farmers depends on both the quality and the market situation.

¹¹For some products, price was missing when using Fenaco data. Yet, these missing values largely concerned the "Others" category of pesticides, and we only use price data for the categories "Herbicides", "Fungicides" and "Insecticides". Plus, as we calculated a mean value over these three categories, we expect the impact of the scarce missing values to be marginal.

¹²Source: Agristat, Statistische Erhebungen und Schätzungen über Landwirtschaft und Ernährung. Swiss Farmers Union. Brugg, Switzerland.

2.6 Insights from the technology choice and the input uses equations

Before presenting the results from the production functions accounting for both technology self-selection and input endogeneity, we present the estimation results we obtained for the technology choice and input use models.

TABLE 2.2: Mean estimates of the coefficients from the Probit model for technology choice

	<i>Adopting low-input practices:</i>
Constant	65.332***
Price ratios	
Herbicide/Wheat	383.684***
Fungicide/Wheat	-397.322***
Insecticide/Wheat	63.730**
Nitrogen/Wheat	-11.982***
Growth regulator/Wheat	-2.997
(Herbicide/Wheat) ²	-56.718***
(Fungicide/Wheat) ²	57.039***
(Insecticide/Wheat) ²	-11.570*
(Nitrogen/Wheat) ²	0.607***
(Growth regulator/Wheat) ²	5.114
Control variables	
French	-0.213
Mountain	0.373**
Wheat share	-0.012*
Farm size	-0.0002***
Weather variables	
log(Rainfall)	-0.573
Temperature	0.200**
Extra variables	
Age	-0.001
In training / apprenticeship	2.706
Training / apprenticeship	-1.782***
Continuing education	-1.468***
Technical schooling	-0.256
Lagged technology change (%)	-0.208***
Annual mean size of low-input farms	-0.014**
Observations	617
AIC	728.219

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

Source: Authors' calculations on Agroscope data.

Table 2.2 presents the estimation results of the technology choice model that considers low- and high-input wheat production practices. As expected from the descriptive statistics on the low- and high-input sub-samples (*cf.* Table 2.1) being located in a mountainous region encourages the adoption of low-input practices whereas farm size and wheat share in farm acreage impact negatively this adoption.

Most price ratios seem to significantly impact the adoption of low-input practices.¹³ In particular, high herbicide and insecticide prices (relatively to wheat prices) are lever to low-input practice adoption, but at a decreasing rate (*cf.* the negative signs of the quadratic terms). On the other hand, high nitrogen and fungicide relative prices are negatively associated to low-input practices adoption (still at a decreasing rate). These last results are quite surprising. Increases in the prices of fungicides or fertilizers are expected to favor the adoption of low-input practices, through a relative profitability effect. Comparison of mean fertilizer use among low- and high-input farmers only shows a 5.4 kg/ha difference (*cf.* Table 2.1), implying that fertilizer prices are expected to have a very limited impact on technology choice.

Other control variables – *e.g.*, age, education and the percentage of technology change that happened at time $t - 1$ and the annual mean size of low-input farms – also appear to significantly impact the production practice choices. In particular, schooling types appear to decrease the low-input adoption probability, indicating differences in terms of farmers' technical abilities as well as differentiate access to the information on low-input practices. The negative effect associated to technology change and the annual mean size of low-input farms are difficult to interpret. Indeed, apart from the price ratios and the weather variables, they are the only variables varying across time. Thus, they might capture year or time trend effects.¹⁴

Parameters estimates of the pesticide and fertilizer use equations for low- and high-input farms are displayed in Tables 2.3 and 2.14. Except for the high-input herbicide input use model, price ratios do not seem to affect input use. Fisher tests to check for the global significance of price ratios in the input use models heads toward the same direction as p-values are superior to 95%, *i.e.* price ratios are not statistically significant in the input use levels models. Importantly, these results do not necessarily demonstrate that input or crop prices don't impact farmers input choices (holding production practices fixed). But, they show that the information content of

¹³A Fisher test on all the price ratios parameters was performed to ensure that our instruments are significant in our technology choice model. The p-value associated to the test was inferior to 1%, *i.e.* relative prices have a significant impact on low-input production practice adoption.

¹⁴We actually check on that point. When integrating year dummies in the technology choice model, the technology change and the mean size of low-input farms are dropped from the model because of singularities. We rather consider the model without these year dummies because these dummies are also absorbing the price ratios effects. Hence, our model divides the time effect in two parts: the price ratio part and a time trend.

the available data don't enable us to accurately estimate the effects of prices on input uses. In particular, prices display limited variability across farms and years in our dataset. Yet, this limited impact of prices on input uses tends to be confirmed by the existing literature. For instance, Skevas, Lansink, and Stefanou (2013) and Böcker and Finger (2017) argue for a price inelasticity of pesticide use.

TABLE 2.3: Results from detailed pesticide uses models for low- and high-input production practices

	<i>Herbicide</i>		<i>Fungicide</i>	<i>Insecticide</i>
	<i>Low-input</i>	<i>High-input</i>	<i>High-input</i>	<i>High-input</i>
Constant	3.559	-8.344	-0.968	-5.238*
Price ratios				
Herbicide/Wheat	-0.311	-13.182	-9.546	-16.572
Fungicide/Wheat	.	21.399	10.274	19.850
Insecticide/Wheat	.	-2.937	2.798	-2.203
Growth regulator/Wheat	.	11.497**	-7.024	0.051
Nitrogen/Wheat	0.130	-0.347	0.574	0.762
(Herbicide/Wheat) ²	3.529	13.028	1.794	4.189
(Fungicide/Wheat) ²	.	-5.444	-1.926	-4.854
(Insecticide/Wheat) ²	.	1.199	-0.972	0.805
(Growth regulator/Wheat) ²	.	-8.179**	3.013	-0.303
(Nitrogen/Wheat) ²	0.003	0.048	-0.027	-0.061
Control variables				
French	0.025	0.095	-0.056	0.016
Mountain	0.164**	-0.009	-0.096	-0.094**
Wheat share	-0.005	0.004	0.018***	0.0005
Farm size	-0.0001**	-0.0002	0.0001**	0.00002
Weather variables				
log(Rainfall)	-0.326*	-0.052	-0.053	0.050
Temperatures	-0.095**	-0.132**	-0.030	0.025
Selection bias				
Inverse Mills ratio	0.038	0.312*	0.420**	0.198**
Observations	381	236	236	236
R ²	0.105	0.170	0.162	0.075
Adjusted R ²	0.079	0.106	0.096	0.003
Residual Std. Error	0.570	0.542	0.589	0.273
	(df = 369)	(df = 218)	(df = 218)	(df = 218)
F Statistic	3.948***	2.633***	2.471***	1.046
	(df = 11; 369)	(df = 17; 218)	(df = 17; 218)	(df = 17; 218)

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

Source: Authors' calculations on Agroscope data.

Another interesting result we get from the input use models is related to selection bias issues. Indeed, inverse Mills ratios are statistically significant in the pesticide use models for high-input farmers. In particular, inverse Mills ratio coefficients in the herbicide, fungicide and insecticide use demand functions of high-input farmers suggest that that high-input farmers have specific unobserved characteristics that tend to increase their pesticide uses. These results are also consistent with self-selection effects.

TABLE 2.4: Results from nitrogen uses model for low- and high-input production practices

	<i>Nitrogen</i>	
	<i>Low-input</i>	<i>High-input</i>
Constant	3.006*	6.493*
Price ratios		
Herbicide/Wheat	-0.587	23.592
Fungicide/Wheat	.	-28.703
Insecticide/Wheat	.	9.657*
Growth regulator/Wheat	.	4.004
Nitrogen/Wheat	0.309	0.038
(Herbicide/Wheat) ²	0.213	-5.714
(Fungicide/Wheat) ²	.	6.605
(Insecticide/Wheat) ²	.	-3.451
(Growth regulator/Wheat) ²	.	-1.837
(Nitrogen/Wheat) ²	-0.036	-0.004
Control variables		
French	0.003	0.027
Mountain	0.048	-0.019
Wheat share	-0.001	0.009***
Farm size	0.00003	-0.00002
Weather variables		
log(Rainfall)	0.204*	-0.433**
Temperatures	0.012	-0.085*
Selection bias		
Inverse Mills ratio	-0.051	0.139
Observations	381	236
R ²	0.028	0.183
Adjusted R ²	-0.001	0.119
Residual Std. Error	0.366	0.375
	(df = 369)	(df = 218)
F Statistic	0.954	2.870***
	(df = 11; 369)	(df = 17; 218)

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

Source: Authors' calculations on Agroscope data.

Other than inverse Mills ratio, control and weather variables best explain pesticide uses. In particular, being located in a mountainous region is associated to an increase in herbicide uses for low-input farmers. Indeed, topographical conditions associated to mountainous areas might make mechanical weeding difficult. On the other hand, being located in a mountainous region tends to be insecticide saving for high-input farmers. Additionally, the wheat acreage share impacts positively herbicide and nitrogen use levels of high-input farmers. A greater wheat acreage share implies greater levels of work if the farmer want to implement mechanical weeding.

In that case, herbicide use tends to be more appealing so that s/he can save time. As for weather conditions, high temperatures have negative impacts on herbicide uses of both low- and high-input farmers as well as negative impacts on nitrogen uses of high-input farmers. Rainfall also affects negatively herbicide use of low-input farmers but affects positively their nitrogen uses. On the contrary, rainfall tends to discourage nitrogen applications among high-input farmers.

Globally, input use demand models are poorly explained by the considered variables as shown by the very small adjusted R^2 . Yet, these models tend to demonstrate that technology selection issues arise when modelling high-input input use levels. Following Di Falco, Veronesi, and Yesuf (2011), we also estimate the effect of the treatment “adopting low-input practices” on the treated (T) – *i.e.* low-input farmers – and on the untreated (U) – *i.e.* high-input farmers – input use levels.

TABLE 2.5: Bootstrapped average treatment effect and the 95% confidence interval on the logarithm of input use levels

Subsample	Input use levels			Treatment effect
	log Herbicide (LI)	log Nitrogen	log Yield	
	-0.092	-0.032	-0.097	Naive
Low-input	-0.028*** [-0.0318; -0.0246]	-0.010*** [-0.0120; 0.0081]	-0.027*** [-0.0292; -0.0258]	TT
High-input	-0.046*** [-0.0550; -0.0356]	-0.016*** [-0.0197; -0.0127]	-0.044*** [-0.0474; -0.0413]	TU

Source: Authors' calculations on Agroscope data.

Note: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. Results were obtained using 1000 bootstrap replications.

We considered the IHS transformation for herbicide use due to zero values.

Table 2.5 presents the estimated treatment effects for input use levels. We also estimated yield supply functions – *i.e.* only price ratios and control variables are considered as covariates – to estimate the average treatment effect on yield levels.¹⁵ First, the estimated treatment effects on the treated and untreated are lower than the naive treatment effect. This confirms the existence of a significant selection bias between low-input adopters and non-adopters among Swiss farmers. In particular, low-input adopters seems to have lower input use and yield levels than high-input farmers due to unobserved characteristics. Accordingly, the naive treatment effect cannot be considered to evaluate the impact of the adoption of low-input production practices on either input uses or yield levels. Additionally, our results show that low-input adoption has a lower impact on low-input input use and yield levels

¹⁵Estimation results for this yield offer functions are available in Appendix 2.9.3.

of low-input farmers than on those of high-input farmers. According to our results, low-input adoption reduces yields of low-input adopters by 2.7% on average whereas this reduction reaches almost 4.5% for high-input farmers.

First and foremost, this section tends to confirm the presence of a production practice selection bias on input uses and yield levels. This encourages us into considering production practice specific input use demand as well as yield supply models. Yet, another important result of this section is that prices do not provide reliable instruments so to explain input use levels. It is important because, undoubtedly, such instruments' weakness reverberates on our attempt to account for input use endogeneity when estimating the production functions.¹⁶

2.7 Production function estimation results

Table 2.6 presents the estimation results of various production function specifications.¹⁷ Model (1) assumes that the production technology is homogeneous, except for a dummy variable indicating the use of low- versus high-input practices (in the potential yield function and in the damage abatement function). No endogeneity issue is considered in this model. Model (2) considers two separate functions for low- and high-input farmers but do not account for any endogeneity issues. Finally, Model (3) corresponds to the endogenous regime switching model accounting for both technology selection bias and input endogeneity issues.

As for the productive part, fertilizers play a significant role in increasing the potential yield in Models (1) and (3). The fact that fertilizer use does not stand out when we consider separate functions while input endogeneity is not accounted for – *i.e.* Model (2) – but does when considering the ERS specification argues in favor of this last specification – *i.e.* Model (3). Plus, when considering the ERS specification, we show that fertilizers play a more important role in the production function of low-input farmers than in the production function of high-input farmers. As for work time, its impact on the potential yield stands out in every considered specification. Yet, Model (3) tends to show that the productive effect of labour is overestimated in the specifications where input endogeneity is not accounted for. Yet, as in Model (2), we find that the productive role of labour between low- and high-input production functions is comparable.

¹⁶For the input use models, we consider the pesticide load indicator. We have also estimated the input use demand models when considering the treatment frequency index as pesticide indicator, in order to see if the model performs better. Overall, estimation results are similar, *e.g.*, price ratios do not appear significant. Thus, the remark on the weakness of instruments still holds.

¹⁷The results we present thereafter consider the load index for pesticide indicators. Results when considering the pesticide treatment frequency indicator (TFI) can be found in Appendix 2.9.4.

TABLE 2.6: Results from different production function specifications accounting for production practices

	(1) All	(2) Low-input High-input	(3) Low-input High-input		
Productive part					
Fertilizers	0.10**	0.06	0.08	0.94***	0.36***
Work & machinery	0.38***	0.33***	0.35***	0.21***	0.20**
French	0.03**	0.01	0.08***	-0.01	0.02
Mountain	-0.01	-0.01	-0.04*	-0.05**	-0.05**
Winter wheat share	0.001	0.002	-0.001	-0.0003	-0.001
Farm size	-0.00001	-0.000003	-0.00001	-0.00004***	0.000001
log(Rainfall)	0.10**	0.16***	0.13**	-0.29***	0.07
Temperatures	0.03***	0.04***	0.04***	0.005	0.04**
Production practice dummy	0.33**
Damage abating part					
Herbicides	0.002	0.06**	-0.004	0.05*	0.04
Insecticides	-0.10	.	-0.11	.	-0.14
Fungicides	0.01	.	0.03	.	-0.05
Mechanical pest control	-0.0003	-0.001	-0.001	-0.0005	-0.0004
Fertilizers	-0.001	-0.0003	-0.001	0.002	0.0005
Constant	-0.19	-0.16	-0.04	-0.43	-0.12
Production practice dummy	-0.32
Decomposition of the error term					
Herbicide endogeneity	.	.	.	-0.05	-0.01
Fungicide endogeneity	0.11
Insecticide endogeneity	0.18
Nitrogen endogeneity	.	.	.	-1.07***	-0.34***
Inverse Mills ratio	.	.	.	0.25**	0.27
Specification					
Separate functions	No	Yes	Yes	Yes	Yes
Endogeneity correction	No	No	No	Yes	Yes
Observations	617	381	236	381	236
R ²	0.998	0.998	0.999	0.998	0.999
Residual Std. Error	0.169	0.180	0.162	0.162	0.153

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

Source: Authors' calculations on Agroscope data.

Regarding the control variable, the ERS specification still tends to stand out from the first two specifications. The negative impact of being located in a mountainous area on the potential yield level that appears among the high-input farmers in Model (2) is generalized to all farmers in Model (3). As for the positive impact of belonging to a French canton on the high-input yield showed by Model (2), it disappears when considering the ERS specification. Moreover, as for the general positive impact of rainfall on wheat yield levels that is revealed in Models (1) and (2), it becomes negative among low-input farmers in Model (3) and loses its significance among high-input wheat producers. However, temperature seems to positively impact these high-input yields, an effect that is not observed in the other specifications. We also notice that, in Model (1), the dummy associated to the production practice suggests that adopting the low-input production practice increases potential yield levels.

As for the damage abatement function, most coefficients are not significant, whatever the considered specification. The only exception is the herbicide coefficient for the low-input production function in both Models (2) and (3). Additionally, both specifications show a similar effect of herbicide in the low-input damage abatement function. Such global lack of significance of the damage abatement function for all considered specifications might come from the input endogeneity that we fail to adequately control for in our ERS model. More globally, the fact that pesticide use levels do not stand out in the yield model may render difficult the study of yield response to pesticide use. Despite that, our ERS framework shows that controlling for input endogeneity – even if poorly – and technology selection issues affect the productive part of the yield model. The significant impact of the input endogeneity and/or selection control functions on the production function tends to confirm this view. In particular, controlling for nitrogen endogeneity seems warranted in both the low- and high-input production functions. Selection issues also seem to impact the low-input production function. Accordingly, unobserved characteristics of low-input farmers tend to increase the observed yield level. It goes in line with the idea of self-selection among low-input farmers.

To see to what extent the changes observed between the three specifications are “yield” impacting, we try and estimate the yield in each sub-sample. Additionally, to see how each production function responds to diminution in pesticide uses, we also estimate yield levels for a 5% and 10% decrease in every pesticide use (*i.e.*, for low-input farmers, we simulate a 5% and 10% decrease in herbicide use and for high-input farmers, we simulate a 5% and 10% decrease in herbicide, fungicide and insecticide uses). Once again, results from this simulation work should be interpreted with caution as our estimated production functions are questionable due to our inability to adequately control for input endogeneity issues.¹⁸

Table 2.7 shows differences in mean yield levels between models. Using separate functions rather than a dummy variable for production practices – Model (1) *vs.* Model (2) – is associated to a decrease in estimated yield levels for both production practices. On the contrary, the model accounting for input endogeneity and selection issues – Model (3) – shows an increase in estimated yield levels for both production practices. Not accounting for input endogeneity and selection issues seems to lead to underestimation of the low-input and high-input yield levels. The yield levels response to the 5% and 10% pesticide decrease scenarios are rather limited for all specifications. Unsurprisingly, overlapping 95% confidence intervals show that these differences are not statistically significant for any of the considered specifications.

¹⁸This cautiousness is even more important given that we did not yet compute the associated confidence intervals.

TABLE 2.7: Bootstrapped estimated yield levels (q/ha) and its 95% confidence interval for the Dummy, Separate and ERS specifications

Specification	Baseline	Pesticide scenarios	
		-5%	-10%
Low-input			
Dummies for production practices	53.13 [52.79; 53.54]	53.15 [52.77; 53.52]	53.15 [52.74; 53.51]
Separate models	51.45 [50.72; 52.29]	51.41 [50.58; 52.21]	51.32 [50.50; 52.12]
ERS models	58.00 [57.62; 58.42]	57.98 [57.56; 58.39]	57.95 [57.53; 58.34]
High-input			
Dummies for production practices	57.72 [57.02; 58.49]	57.70 [56.97; 58.40]	57.64 [56.96; 58.36]
Separate models	54.99 [53.91; 55.98]	54.88 [53.86; 55.94]	54.83 [53.78; 55.83]
ERS models	65.36 [64.47; 66.15]	65.49 [64.59; 66.26]	65.58 [64.79; 66.39]

Source: Authors' calculations on Agroscope data.

Note: Results were obtained using 1000 bootstrap replications.

When considering the observed mean yield levels among each sub-samples (Table 2.1), the ERS model seems to yield more reasonable results for both the high-input and low-input production functions in comparison to the other models. Indeed, the empirical mean yield level of the high-input farmers, respectively low-input farmers, is around 64 q/ha, respectively 58 q/ha.¹⁹ Still, we need to remain cautious when interpreting such estimation results for two main reasons. First, the sub-sample empirical means do not account for selection biases. Second, apart from the weak-instruments issue, estimated mean yield levels are very sensitive to the chosen specification.

Robustness checks

In order to evaluate the sensitivity of our ERS specification, we performed six different robustness checks. Robustness check (1a), (1b) and (2) correspond to the robustness checks associated to the asymmetry and separability tests (*cf.* Appendix 2.9.1). In particular, robustness checks (1a) and (1b) consider respectively work and machinery as interactive input and no interactive input at all. Robustness check (2) corresponds to the model without damage abatement function, *i.e.* we consider a Cobb-Douglas production function where all inputs are assumed to play a symmetric role. As

¹⁹From that viewpoint, results obtained when considering the pesticide treatment frequency indicator (TFI) rather than the pesticide load index (LI) seem less satisfactory (see Appendix 2.9.4). Indeed, in particular for low-input farmers, the estimated yield level is quite superior to what is observed in the sub-sample.

for robustness check (3), we consider the global pesticide variable instead of the separate levels of herbicide, insecticide and fungicide. Robustness check (4) considers a Translog production function instead of the Cobb-Douglas. Finally, robustness check (5) considers revenue as the explanatory variable. The estimated coefficients for each robustness checks can be found in Appendix (*cf.* Appendices 2.9.1, 2.9.5, 2.9.7 and 2.9.6).

TABLE 2.8: Bootstrapped estimated yield levels (q/ha), or revenue (CHF/ha), and its 95% confidence interval obtained for the different robustness checks

Specification	<i>Baseline</i>	<i>Pesticide Scenarios</i>	
		-5%	-10%
Low-input			
Main model	58.00	57.98	57.95
	[57.62; 58.42]	[57.56; 58.39]	[57.53; 58.34]
Robustness check (1a)	56.26	56.14	56.02
	[55.90; 56.63]	[55.79; 56.51]	[55.66; 56.38]
Robustness check (1b)	63.99	63.97	63.95
	[63.55; 64.46]	[63.55; 64.41]	[63.51; 64.39]
Robustness check (2)	55.68	55.33	54.96
	[55.33; 56.04]	[54.96; 55.70]	[54.62; 55.32]
Robustness check (3)	54.56	54.52	54.44
	[54.16; 54.99]	[54.08; 54.93]	[54.04; 54.89]
Robustness check (4)	46.41	46.34	46.33
	[45.66; 47.27]	[45.60; 47.25]	[45.50; 47.13]
Robustness check (5)	3176.60	3177.20	3176.77
	[3143; 3210]	[3142; 3210]	[3144; 3209]
High-input			
Main model	65.36	65.49	65.58
	[64.47; 66.15]	[64.59; 66.26]	[64.79; 66.39]
Robustness check (1a)	66.63	66.74	66.84
	[65.82; 67.43]	[65.89; 67.54]	[66.05; 67.59]
Robustness check (1b)	66.01	66.00	65.98
	[65.14; 66.83]	[65.17; 66.79]	[65.20; 66.78]
Robustness check (2)	64.35	64.09	63.86
	[63.57; 65.03]	[63.35; 64.84]	[63.12; 64.58]
Robustness check (3)	68.09	67.92	67.80
	[67.31; 68.86]	[67.22; 68.74]	[67.09; 68.50]
Robustness check (4)	83.80	84.73	85.61
	[81.99; 85.53]	[82.84; 86.41]	[83.79; 87.34]
Robustness check (5)	2813.61	2813.55	2812.47
	[2755; 2878]	[2754; 2877]	[2756; 2874]

Source: Authors' calculations on Agroscope data.

Note: Results were obtained using 1000 bootstrap replications. For robustness check (5), the unit considered is not q/ha but CHF/ha as we consider revenue as the variable to explain.

The most striking result when looking at Table 2.8 is the great variability of the estimated yield levels upon the chosen specification. Estimated low-input yield levels, respectively high-input yield levels, at baseline vary from 46.41 q/ha to 63.99 q/ha, respectively from 64.35 q/ha to 83.80 q/ha. More specifically, robustness check (4) – *i.e.* the one with the Translog production function – is the robustness check associated with the greatest changes in estimated yield levels. For the others robustness checks, the differences with the benchmark specification are less marked. Yet, as for the benchmark specification, robustness checks display a systematic higher estimated yield level for high input farmers compared to low-input ones. This feature is shared by all specifications, as well as the non-significant yield differences when cutting pesticide uses by 5% or 10%. When considering revenue as the variable to be modelled, our results tend to show a greater revenue for low-input farmers compared to high-input farmers, which is in line with the descriptive statistics on the sub-samples.

The variability in estimated yield levels should not hide the robustness of some results across our specifications.²⁰ First, the larger effects of fertilizers in the low-input production function in comparison to the high-input one stands out from most specifications. Similarly, we find that the role of labour on yield levels is similar across production practices. More importantly, the results on input endogeneity and selection biases are also consistent across (most) specifications. In particular, the nitrogen endogeneity negatively affects all low- and high-input production functions. We also find that the selection bias impacts positively the low-input production function across specifications (except when we consider revenue rather than yield as the variable to explain).

Our estimates the technology specific production functions clearly suffer from identification issues. This is due to our lacking suitable instrumental variables for instrumenting input use levels. This prevents us from deriving robust conclusions from our results. Nonetheless, the estimation results we obtain for the input demand and yield supply models strongly suggest that one cannot assess the effects of the adoption of the low-input practices without considering selection issues. In other words, important unobserved drivers of the yield and input use levels of Swiss wheat producers also impact their production practices choice. Neglecting the effects of these unobserved drivers tends to significantly bias the comparison of the performances of the considered production technologies.

²⁰The following results applied for all robustness check except the Translog one. Indeed, when considering the Translog function form for the productive part of the function, we struggle to find significant effects (cf. Appendix 2.9.7).

2.8 Discussion and Conclusion

2.8.1 The input endogeneity issue

As said previously, our ERS framework suffers from weak instruments to control for input endogeneity. For now, as calculated, pesticide prices vary only across year. Yet, we only have 7 years in our panel dataset, *i.e.* the variation of the price ratio variables is rather limited. This issue is with econometric production functions estimated with short panel data. Instrument variables other than price variables are difficult to uncover. Farmers adapt their input uses to exogenous factors when these factors impact their crop yield levels, implying that the exclusion restriction necessary for characterizing instrumental variables does not hold.

Another solution that can be investigated is to consider a full reduced form for both the input use demand and the yield demand models. Even if does not allow to uncover for the yield response to input use variations, full reduced form specification permits to estimate the difference between low- and high-input yield and input use levels while accounting for the potential selection issues.

2.8.2 General considerations

Either considering a primal or dual optimization problem, an interesting lead so to improve the framework we consider in this article is to introduce a dynamic dimension for the production practices choice. The production practices adopted at time $t - 1$ might impact the production practices choice at time t but also input use levels and achieved yield. In particular, crop rotation is often well-valued among low-input practices. Crop rotation can affect the input use levels in the future – *e.g.*, it decreases the need for chemical protection – and thus might affect yield as well. One could investigate such time-interdependence either by allowing an heteroskedastic form of the error term or by introducing lagged variables in the production and input use functions.

2.8.3 Conclusion

Despite our attempt to control for input endogeneity, we only have weak instruments, *i.e.* input endogeneity control functions in the yield level model are not expected to correct for input endogeneity. The limited effect of prices on input uses is well-known among economists (see, *e.g.*, Böcker and Finger, 2017; Skevas, Lansink, and Stefanou, 2013).

Still, even if not accounting correctly for the input endogeneity issue, the ERS specification shows differences in comparison to the specification where technology selection bias is not accounted for. In particular, it shows a larger productive role of fertilizers among low-input farmers. Additionally, estimation results from both the

production and input use demand models tend to argue in favor of a selection bias. As for high-input farmers, this selection bias affects positively their input uses, *i.e.* their unobserved characteristics force them into use greater input use levels. On the other side, low-input farmers show a positive selection bias in the yield model, *i.e.* their unobserved characteristics allow them to obtain greater yield levels than the average when using low-input production practices. Our results also tend to show that the naive estimation of the treatment effect is not adequate to estimate the impact of low-input adoption on both input use and yield levels. All results combined are very interesting to understand the low-input adoption process. In that sense our findings, even if they have to be considered very cautiously, confirm the findings from Finger and El Benni (2013) on the characteristics of the low-input adopters and more globally the adverse selection issue pointed out by Pietola and Lansink (2001) for AEP adopters.

Overall, we can consider Swiss wheat as an example of a successful diffusion of low-input practices with more than 50% of adoption rate. Further research could consider performing a cost-benefit analysis to see how much the diffusion of low-practices cost to the government and what are the benefit in terms of pesticide reduction. In particular, as the high-input farmers are farmers with the most intensive use of pesticide, the impact on such policy on pesticide reduction might be rather limited in comparison to the deployed efforts. Another arising question concerns the diffusion dynamics of these low-input practices. As said before, low-input practices are successfully adopted by a large share of wheat farmers. Yet, as the remaining high-input farmers tend to have unobserved specific characteristics boosting up their pesticide use, we could wonder how to incentivize them to adopt these low-input practices. In particular, drawing a parallel between the low-input diffusion among Swiss wheat farmers and the technology diffusion theory of Rogers (1962), the remaining high-input farmers could be considered as “laggards” that remained to be convince to adopt the low-input technology.

2.9 Appendices

2.9.1 Results from the separability and asymmetry tests

Results from separability test and associated robustness check

Separability tests are performed in order to determine whether inputs have a purely productive or damage abatement function or if they have an interactive effect (see Saha, Shumway, and Havenner, 1997). Let $\mathbf{x}_{(d)}$ be the vector of purely damage abating inputs, *i.e.* pesticide uses and mechanical pest control. Let $\mathbf{x}_{(p)}$ be the vector of productive inputs, *i.e.* fertilizers and work. Yet, as mentioned previously, one could think that fertilizers and/or can interact with pesticide uses. To test for their interactive effect, we estimate the productive part of the yield function $h(\cdot)$ with $\mathbf{x}_{(p)}$, $\mathbf{x}_{(d)}$ *i.e.* both productive and protective inputs and their interactions. Thus we can write the following yield function

$$y_{it} = s(\mathbf{c}, \boldsymbol{\beta}_{(c)}) h(\mathbf{x}_{(p)}, \mathbf{x}_{(d)}; \boldsymbol{\beta}_{(p)}, \boldsymbol{\beta}_{(d)}, \boldsymbol{\beta}_{(pd)}), \quad (2.13)$$

where $\boldsymbol{\beta}_{(p)}$ (resp. $\boldsymbol{\beta}_{(d)}$) represents the vector associated to $\mathbf{x}_{(p)}$ (resp. $\mathbf{x}_{(d)}$) and $\boldsymbol{\beta}_{(pd)}$ denotes the vector of the coefficients associated to the interactive terms between. To check for the separability of inputs, we test for the significance of all interactive terms at the same time *i.e.*

$$\begin{cases} H_0 : \boldsymbol{\beta}_{pd} = 0 \\ H_1 : \exists k_{(p)}, \boldsymbol{\beta}_{pd} \neq 0 \end{cases} ,$$

where $k_{(p)}$ corresponds to the index of potentially interactive input (here work and fertilizers). If the hypothesis that all interactive terms are equal to zero (H_0) is rejected then we consider the input with the greatest test statistic. Once removed from H_0 , we check if the remaining interactive variables are equal to 0. If so, the only interactive input is the one we removed. If not, it means in our case that both work and fertilizers can be considered as interactive inputs. the corresponding input will be considered as interactive. To have robust results, different functional form were used to perform those separability tests: a Cobb-Douglas, a quadratic and a Translog functional form.²¹ Results are gathered in Table 2.9.

TABLE 2.9: Results from separability tests

Specification	All farms	Low-input	High-input
Cobb-Douglas	work	work & fert	work
Quadratic	∅	fert	∅
Translog	∅	∅	∅

Source: Authors' calculations on Agroscope data.

²¹To avoid 0 issues when considering the Cobb-Douglas and Translog functional forms, we considered the IHS transformation instead of the standard logarithmic to linearize the yield function (see, Bellemare, Barrett, and Just, 2013 for instance).

When considering all farms, the only variable that seems to have a potential interactive effect is work. Results are similar for high-input production practices farmers. Yet, when considering low-input production practices farmers, separability tests tend to show that fertilizers should be considered as an interactive input. As the main analysis was performed considering fertilizers as an interactive input, we present thereafter the result when considering (i) no interactive inputs in both production practices as well as (ii) work and fertilizers as interactive inputs for low-input farmers and work as an interactive input for high-input farmers.

TABLE 2.10: Asymmetry robustness check - Production function estimates with different interactive inputs

	<i>Low-input</i>		<i>High-input</i>	
	(i)	(ii)	(i)	(ii)
Productive part				
Fertilizers	0.98***	0.84***	0.36***	0.32***
Work & machinery	0.21***	0.28***	0.21***	0.30**
French	-0.02	-0.02	0.03	0.01
Mountain	-0.06**	-0.05**	-0.05*	-0.06**
Winter wheat share	0.0004	-0.00004*	-0.002	-0.0006
Farm size	-0.00003***	-0.00003***	-0.000004	0.000007
log(Rainfall)	-0.28***	-0.28***	0.06	0.02
Temperatures	0.007	0.004	0.04**	0.04**
Damage abating part				
Herbicides	0.07**	-0.07**	0.01	-0.08*
Insecticides	.	.	-0.14	0.11
Fungicides	.	.	-0.03	0.02
Mechanical pest control	-0.0002	0.0002	-0.0004*	0.0001
Fertilizers	.	-0.002	.	.
Work & machinery	.	0.0001	.	0.0001
Constant	-0.63*	0.40	-0.04	0.15
Decomposition of the error term				
Herbicide endogeneity	-0.13**	-0.07	0.01	-0.10
Fungicide endogeneity	.	.	0.10	0.08
Insecticide endogeneity	.	.	0.22	0.13
Nitrogen endogeneity	-0.99***	-1.001***	-0.32***	-0.31***
Inverse Mills ratio	0.22**	0.24**	0.33	0.02
Observations		381		236
R ²	0.998	0.998	0.999	0.999
Residual Std. Error	0.163	0.162	0.153	0.153

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

Source: Authors' calculations on Agroscope data.

Results from asymmetry test and associated robustness check

Asymmetry tests are aimed to see whether or not it is relevant to consider that inputs have an asymmetric role on the crop production (*i.e.* we need to separate the productive and damage abatement parts) or if their role is symmetric (*i.e.* we just need to estimate the productive part with all inputs) (see, Zhengfei et al., 2006).²² To do so, we consider both the productive and damage abating part of the production function and assign each input to its corresponding part. Then, purely damage abating inputs are included in the productive part and associated coefficients are tested to zero. At the same time, we test whether or not coefficients in the damage abating part are simultaneously equal to zero. Let $\mathbf{x}_{(p/i)}$ (resp. $\mathbf{x}_{(d/i)}$) represents the set of purely productive (resp. protective) inputs and $\mathbf{x}_{(i)}$ the set of interactive inputs. We thus have

$$y_{it} = s(\mathbf{c}, \boldsymbol{\beta}_{(c)}) h(\mathbf{x}_{(p/i)}, \mathbf{x}_{(i)}, \mathbf{x}_{(d/i)}; \boldsymbol{\beta}_{(p/i)}, \boldsymbol{\beta}_{(i)}^p, \boldsymbol{\beta}_{(d/i)}^p) g(\mathbf{x}_{(d/i)}, \mathbf{x}_{(i)}; \boldsymbol{\beta}_{(d/i)}^d, \boldsymbol{\beta}_{(i)}^d), \quad (2.14)$$

where $\boldsymbol{\beta}_{(d/i)}^p$ (resp. $\boldsymbol{\beta}_{(d/i)}^d$) represents the vector of parameters associated to the purely protective inputs in the productive (resp. protective) part of the yield function. Then, the two tests that are performed are

$$\begin{cases} H_0 : \boldsymbol{\beta}_{(d/i)}^p = 0 \\ H_1 : \exists k_{(d/i)}, \boldsymbol{\beta}_{(d/i)}^p \neq 0 \end{cases} \quad ,$$

and

$$\begin{cases} H_0 : \boldsymbol{\beta}_{(d/i)}^d = \boldsymbol{\beta}_{(i)}^d = 0 \\ H_1 : \exists (k_{(d/i)}, k_{(i)}), (\boldsymbol{\beta}_{(d/i)}^p, \boldsymbol{\beta}_{(i)}^p) \neq 0 \end{cases} \quad .$$

If we accept the first test and reject the second test then we can conclude that there is asymmetry among inputs. On the contrary, if we reject the first test and accept the second one, it means that inputs play a symmetrical role in the production function. Similarly to what we have done with separability tests, we used several specification for the productive functional form (*e.g.*, Cobb-Douglas, quadratic and Translog). Results are gathered in Table 2.11. We emphasized the cases where asymmetry was confirmed (*i.e.* cases where we have $\{H_0, H_1\}$).

Whatever the chosen specification, those asymmetry tests tend to show that the asymmetry hypothesis is poorly verified. The only cases when asymmetry seems to hold is when we consider a Translog function and (*i*) when we consider work as interactive input or (*ii*) when we consider only the low-input production practices farms.

²²Asymmetry test is to be performed after the separability test to ensure a correct allocation of inputs to their respective role.

TABLE 2.11: Results from the asymmetry test

Specification	All farms		
	Fert	Work	\emptyset
Cobb-Douglas	$\{H_1, H_1\}$	$\{H_1, H_1\}$	$\{H_1, H_1\}$
Quadratic	$\{H_1, H_1\}$	$\{H_1, H_1\}$	$\{H_1, H_1\}$
Translog	$\{H_1, H_1\}$	$\{H_0, H_1\}$	$\{H_1, H_1\}$

Specification	Low-input			High-input		
	Fert	Fert & Work	\emptyset	Fert	Work	\emptyset
Cobb-Douglas	$\{H_1, H_1\}$	$\{H_1, H_1\}$	$\{H_1, H_1\}$	$\{H_1, H_1\}$	$\{H_1, H_1\}$	$\{H_1, H_1\}$
Translog	$\{H_0, H_1\}$	$\{H_0, H_1\}$	$\{H_0, H_1\}$	$\{H_1, H_1\}$	$\{H_0, H_1\}$	$\{H_1, H_1\}$

Source: Authors' calculations on Agroscope data.

In line with these results, we perform a robustness check considering a symmetrical role for inputs and using a Cobb-Douglas production function *i.e.*

$$y_{it} = s(\mathbf{c}, \boldsymbol{\beta}_{(c)}) h(\mathbf{x}; \boldsymbol{\beta}) = \exp\left(\sum_c \beta_{(c)} c\right) \prod_x x^\beta$$

where \mathbf{c} represents the vector of control variable and \mathbf{x} the vector of inputs. Results are presented thereafter in Table 2.12.

When using a yield function with no damage abatement part, results tend to be similar for both production practices and endogeneity issues are not relevant anymore in our model. Hence, a correct specification of the production function is critical if we want to account for the effect of technology choice on production function.

TABLE 2.12: Symmetry robustness check - Production function estimates with symmetry of inputs

	<i>Low-input</i>	<i>High-input</i>
Cobb-Douglas production function		
log Fertilizers	1.03***	0.51***
log Work & machinery	0.09	0.12
log Herbicides	0.31***	0.21**
log Insecticides	.	-0.18
log Fungicides	.	-0.03
log Mechanical pest control	-0.02***	-0.006
French	-0.04*	-0.006
Mountain	-0.09***	-0.06
Winter wheat share	0.002*	-0.002
Farm size	-0.00001	0.00002
log(Rainfall)	-0.28***	0.14***
Temperatures	0.01	0.03*
Decomposition of the error term		
Herbicide endogeneity	-0.31***	-0.25**
Fungicide endogeneity	.	0.02
Insecticide endogeneity	.	-0.11
Nitrogen endogeneity	-0.99***	-0.55***
Inverse Mills ratio	0.12	-0.36
Observations	381	236
R ²	0.999	0.999
Residual Std. Error	0.164	0.167

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

Source: Authors' calculations on Agroscope data.

2.9.2 Additional descriptive statistics

In Table 2.13 we present the frequency of observations and the sample composition between Extenso and Intenso from 2009–2015. Intenso farms are always less represented than Extenso ones and this phenomenon increases over years.

TABLE 2.13: Repartition of Low-input / High-input farms among years

	<i>Low-input</i> N (%)	<i>High-input</i> N (%)
2009	65 (53%)	57 (47%)
2010	59 (59%)	41 (41%)
2011	53 (54%)	45 (46%)
2012	60 (68%)	28 (32%)
2013	51 (65%)	27 (35%)
2014	47 (65%)	25 (35%)
2015	46 (78%)	13 (22%)
Total	381	236

Source: Authors' calculations
on Agroscope data.

2.9.3 Estimation results from yield offer functions

TABLE 2.14: Results from yield offer model for low- and high-input production practices

	<i>Yield</i>	
	<i>Low-input</i>	<i>High-input</i>
Constant	4.638***	7.348***
Price ratios		
Herbicide/Wheat	0.582	8.694
Fungicide/Wheat	.	-11.775
Insecticide/Wheat	.	2.585
Nitrogen/Wheat	0.341**	0.261
Growth regulator/Wheat	.	0.202
(Herbicide/Wheat) ²	-0.196	-1.703
(Fungicide/Wheat) ²	.	2.379
(Insecticide/Wheat) ²	.	-0.882
(Nitrogen/Wheat) ²	-0.029**	-0.030
(Growth regulator/Wheat) ²	.	0.220
Control variables		
French	-0.048**	-0.030
Mountain	0.027	-0.018
Wheat share	0.001	0.003**
Farm size	-0.00000	-0.00002*
Weather variables		
log(Rainfall)	-0.266***	-0.278***
Temperatures	-0.016	-0.020
Selection bias		
Inverse Mill's ratio	0.038	-0.163***
Observations	381	236
R ²	0.100	0.162
Adjusted R ²	0.073	0.096
Residual Std. Error	0.162	0.172
	(df = 369)	(df = 218)
F Statistic	3.739***	2.474***
	(df = 11; 369)	(df = 17; 218)

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

Source: Authors' calculations on Agroscope data.

2.9.4 Estimation results when considering the TFI pesticide indicator

TABLE 2.15: Results from detailed pesticide uses models for low- and high-input production practices, with TFI pesticide indicator

	<i>Herbicide</i>		<i>Fungicide</i>	<i>Insecticide</i>
	<i>Low-input</i>	<i>High-input</i>	<i>High-input</i>	<i>High-input</i>
Constant	1.088	-0.940	4.296	-0.366
Price ratios				
Herbicide/Wheat	0.532	49.568*	-20.063	-14.840
Fungicide/Wheat	.	-49.288*	19.196	16.243
Insecticide/Wheat	.	6.392	-0.190	-1.824
Growth regulator/Wheat	.	11.364**	-8.534*	-4.728*
Nitrogen/Wheat	0.126	-1.994**	1.091	0.625
(Herbicide/Wheat) ²	-0.169	-12.704*	3.763	3.270
(Fungicide/Wheat) ²	.	12.154*	-3.440	-3.478
(Insecticide/Wheat) ²	.	-2.292	-0.283	0.512
(Growth regulator/Wheat) ²	.	-7.342***	4.904	3.121**
(Nitrogen/Wheat) ²	-0.005	0.177**	-0.078	-0.052
Control variables				
French	-0.128***	0.058	-0.135*	-0.099**
Mountain	0.098**	0.077	-0.082	-0.038
Wheat share	0.002	0.005	0.015***	0.002
Farm size	-0.00002	-0.00004	0.00005	-0.00000
Weather variables				
log(Rainfall)	-0.095	0.069	-0.407*	-0.138
Temperatures	-0.027	-0.024	-0.003	0.073***
Selection bias				
Inverse Mills ratio	-0.030	0.068	0.076	0.090
Observations	381	236	236	236
R ²	0.081	0.144	0.248	0.201
Adjusted R ²	0.054	0.077	0.189	0.139
Residual Std. Error	0.328	0.420	0.459	0.234
F Statistic	(df = 369) 2.975***	(df = 218) 2.157***	(df = 218) 4.223***	(df = 218) 3.228***
	(df = 11; 369)	(df = 17; 218)	(df = 17; 218)	(df = 17; 218)

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

Source: Authors' calculations on Agroscope data.

Table 2.15 seems to indicate when considering TFI pesticide indicator rather than LI, some price ratios gain significance in the input demand models, in particular for the high-input herbicide demand model. Yet, when realizing the Fisher test for the global significance of the price instruments, we still reject the null hypothesis, *i.e.* price ratios are not statistically significant in our models. Plus, coefficients associated to the inverse Mills ratio lose their significance, *i.e.* selection issue does not stand out in this specification. We also performed the production function estimation with the TFI pesticide indicator in order to see in what extent results change when considering another pesticide indicator. Results are presented in Tables 2.16 and 2.17.

TABLE 2.16: Results from different production function specifications accounting for production practices, with TFI indicator for pesticide use

	(1)	(2)		(3)	
	All	Low-input	High-input	Low-input	High-input
Productive part					
Fertilizers	0.10**	0.12*	0.09*	1.01***	0.53***
Work & machinery	0.37***	0.42***	0.23***	0.21***	0.10
French	0.04**	0.005	0.05	0.04	-0.04
Mountain	-0.01	0.01	-0.03	-0.08**	-0.01
Winter wheat share	0.001	0.001	-0.0001	-0.001	-0.002*
Farm size	-0.00001	-0.00001	-0.00001	-0.00003***	-0.000004
log(Rainfall)	0.11***	0.08*	0.21***	-0.29***	0.05
Temperatures	0.03***	0.03***	0.07***	0.006	0.06***
Production practices	0.28*
Damage abating part					
Herbicides	0.002	0.02	-0.13	-0.23	-0.16*
Insecticides	0.18*	.	-0.42	.	-0.50*
Fungicides	0.01	.	0.15**	.	0.14**
Mechanical pest control	-0.0003	0.0003***	-0.00001	0.0002	-0.0001
Fertilizers	-0.001	0.001**	-0.002**	-0.0001	-0.0017**
Constant	-0.16	0.38***	0.17	0.84	0.24
Production practices	-0.29
Decomposition of the error term					
Herbicide endogeneity	.	.	.	-0.51*	0.04
Fungicide endogeneity	0.02
Insecticide endogeneity	0.19
Nitrogen endogeneity	.	.	.	-1.05***	-0.43***
Selection bias	.	.	.	0.18	0.21
Specification					
Separate functions	No	Yes	Yes	Yes	Yes
Endogeneity correction	No	No	No	Yes	Yes
Observations	617	381	236	381	236
R ²	0.998	0.998	0.998	0.998	0.998
Residual Std. Error	0.169	0.174	0.182	0.163	0.170

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

Source: Authors' calculations on Agroscope data.

TABLE 2.17: Bootstrapped estimated yield levels and its 95% confidence intervals obtained with our Dummy, Separate and ERS specifications, with TFI indicator for pesticide use

Specification	<i>Baseline</i>	<i>Pesticide scenarios</i>	
		–5%	–10%
Low-input			
Model (1)	46.25 [45.93; 46.56]	46.24 [45.94; 46.54]	46.24 [45.93; 46.52]
Model (2)	56.80 [56.41; 57.21]	56.90 [56.51; 57.31]	57.00 [56.60; 57.42]
Model (3)	63.10 [62.70; 63.50]	62.09 [61.70; 62.46]	61.04 [60.68; 61.42]
High-input			
Model (1)	51.94 [51.16; 52.65]	51.81 [51.09; 52.53]	51.67 [50.97; 52.47]
Model (2)	66.29 [65.43; 67.18]	66.43 [65.61; 67.32]	66.56 [65.82; 67.41]
Model (3)	67.46 [66.70; 68.31]	67.73 [66.92; 68.52]	67.97 [67.19; 68.74]

Source: Authors' calculations on Agroscope data.

Note: Results were obtained using 1000 bootstrap replications.

2.9.5 Results when using global pesticide use variable in the production function

We present here the results from the production function accounting for both self-selection and endogeneity when using a global pesticide variable instead of separate variables for fungicides, insecticides and herbicides.

TABLE 2.18: Global pesticide robustness check - Production function estimates with global pesticide indicator

	<i>Low-input</i>	<i>High-input</i>
Productive part		
Fertilizers	0.93***	0.40***
Work & machinery	0.21***	0.22***
French	-0.01	0.04
Mountain	-0.05*	-0.03
Winter wheat share	-0.0002	-0.004**
Farm size	-0.00004***	-0.00001
log(Rainfall)	-0.28***	0.06
Temperatures	0.005	0.04***
Damage abating part		
Pesticides	0.06**	-0.08*
Mechanical pest control	-0.0005	0.0004
Fertilizers	0.001	-0.0003
Constant	-0.42	0.33
Decomposition of the error term		
Pesticide endogeneity	-0.02	-0.02
Nitrogen endogeneity	-1.05***	-0.41***
Inverse Mills ratio	0.23**	0.10
Observations	381	236
R ²	0.998	0.998
Residual Std. Error	0.162	0.170

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

Source: Authors' calculations on Agroscope data.

2.9.6 Results when using revenue instead of yield in the production function

We present here the results from the production function accounting for both self-selection and endogeneity when using revenue as explained variable instead of yield.

TABLE 2.19: Revenue robustness check - Production function estimates with revenue as the explained variable

	<i>Low-input</i>	<i>High-input</i>
Productive part		
Fertilizers	1.97***	0.90***
Work & machinery	0.21***	0.20*
French	0.01	0.05
Mountain	-0.13***	-0.08**
Winter wheat share	-0.002	-0.004**
Farm size	-0.00006***	0.00005**
log(Rainfall)	-0.38***	0.27***
Temperatures	-0.005	0.09***
Damage abating part		
Herbicides	-0.0004	0.10**
Insecticides	.	-0.10
Fungicides	.	-0.03
Mechanical pest control	-0.0005	-0.0003
Fertilizers	0.002	0.0002
Constant	-0.39	-0.43
Decomposition of the error term		
Herbicide endogeneity	-0.009	-0.20*
Fungicide endogeneity	.	-0.02
Insecticide endogeneity	.	-0.05
Nitrogen endogeneity	-2.04	-1.02***
Inverse Mills ratio	-0.33**	-0.79***
Observations	381	236
R ²	0.999	0.999
Residual Std. Error	0.207	0.196

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

Source: Authors' calculations on Agroscope data.

2.9.7 Results when using a Translog specification for the productive part

We present here the results from the production function accounting for both self-selection and endogeneity when using a Translog specification for the productive part instead of a Cobb-Douglas. Translog functions are known for being more flexible than Cobb-Douglas.

TABLE 2.20: Translog robustness check - Production function estimates with Translog specification for the productive part

	<i>Low-input</i>	<i>High-input</i>
Productive part		
Fertilizers	-0.34	0.34
Fertilizers ²	-0.01	0.09
Work & machinery	1.64***	0.97
Work & machinery ²	-0.27**	-0.08
Fertilizers & Work	-0.01	-0.15
French	-0.03	0.03
Mountain	-0.02	-0.06**
Winter wheat share	0.0009	0.0007
Farm size	-0.00001	-0.00001
log(Rainfall)	-0.23***	-0.12
Temperatures	-0.003	0.004
Damage abating part		
Herbicides	-0.05*	0.04
Insecticides	.	-0.15
Fungicides	.	-0.10**
Mechanical pest control	0.0004	0.00004
Fertilizers	-0.002	0.001
Constant	0.56	-0.06
Decomposition of the error term		
Herbicide endogeneity	-0.05	0.03
Fungicide endogeneity	.	0.08
Insecticide endogeneity	.	0.15
Nitrogen endogeneity	-0.25	-0.08
Inverse Mills ratio	0.15	0.23
Observations	381	236
R ²	0.999	0.999
Residual Std. Error	0.159	0.149

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

Source: Authors' calculations on Agroscope data.

Chapter 3

Crop production models accounting for latent CMP choices ¹

3.1 Introduction

Regulation of pollutions due to the use of chemical inputs,² pesticides in particular, in agricultural production is a major policy objective in the European Union (EU) (see, Directive 2009/128/EC³). However, most regulation policies implemented until now have achieved limited reductions in the use of these polluting inputs, especially regarding pesticides. Within the EU, these policies have been mainly based on market access restrictions, funding public research efforts on chemical fertilization and crop protection alternatives as well as subsidies aimed to disseminate chemical input saving crop production practices. While economists generally advocate for implementing taxes to internalize the negative external effects of input uses, public decision makers are reluctant to use taxes owing to their potential impact on farmers' income.⁴ As a matter of fact, chemical inputs are key production factors in the crop production technologies used by farmers in industrialized countries (*e.g.*, Lin, 2011; Matson et al., 1997; Tilman et al., 2002 and Aubertot et al., 2007 for a focus on France). Production practices targeting high yield levels and based on short crop rotation schemes require high chemical fertilization levels. They generate high potential yield levels, which are worth protecting, and increase pest, disease and weed risks (*e.g.*, Boquet and Johnson, 1987; Howard, Chambers, and Logan, 1994; Roth et al., 1984). As a result, these practices call for high protection levels that can be achieved by farmers relatively easily and at reasonable (private) costs by relying on chemical pesticides.⁵ Hence, the political reluctance to tax these inputs as it is expected to

¹This chapter is written in collaboration with Obafémi Philippe KOUTCHADÉ and Alain CARPENTIER.

²Production and use of mineral nitrogen emit greenhouse gases and excess nitrogen uses pollute water bodies. Mineral phosphorus is a nonrenewable resource with rapidly decreasing stocks that generate eutrophication of surface when used in excess.

³*Directive 2009/128/EC of the European Parliament and of the Council of 21 October 2009 establishing a framework for Community action to achieve the sustainable use of pesticides 2009.*

⁴For instance, pesticide taxes were considered by the European Commission but have not been implemented. In the few countries where they are, pesticide taxes are generally implemented with low tax rates and mostly for fund raising (Finger et al., 2017; Skevas, Lansink, and Stefanou, 2013).

⁵Indeed, conventional crop production practices have been designed while taking for granted that high crop protection levels can be achieved at reasonable (private) cost with chemical pesticides (see,

significantly impact farmers' income with limited impact on their uses and the need for a policy scheme impacting chemical input uses rather than farmers' income.

Econometric results which demonstrate that farmers' pesticide uses generally display very limited responsiveness to pesticide price increases (see, *e.g.*, Böcker and Finger, 2017; Skevas, Lansink, and Stefanou, 2013), seem to provide support to the hypothesis stating that the current agricultural production technology "heavily depends" on pesticide uses and thus justify such political reluctance. Yet, micro-econometric analyses of farmers' chemical input uses are generally based on panel data with short time dimension and assume that farmers' production technology remains unchanged throughout the considered period. Hence, these analyses mostly reveal the inelasticity of farmers' chemical input choices given their current technology choice, a result consistent with the view of agricultural scientists. According to them, farmers cannot significantly modify their chemical input uses without changing their cropping management practices (CMPs) or cropping systems. Properties of CMPs, which consist in ordered sequence of operations aimed to produce a given crop, are only guaranteed to hold in a limited range of chemical input uses (in a given production area). Adopting alternative production practices is necessary for farmers to significantly reduce their use of chemical inputs with limited impact on their income.

Thus our considering of CMP choice in this article. Farmer's CMP choice can be considered as a short run choice that is not affected by huge investment costs. From that perspective, it shares much more similarities with crop variety choice, which is usually considered and modelled as a short run choice (*e.g.*, Michler et al., 2018; Suri, 2011), than with irrigation technology choice, which is modelled as an investment choice (*e.g.*, Genius et al., 2014). Purchasing and installing irrigation equipment entails significant sunk costs. The (switching) costs related to CMP or seed variety changes are rather limited. Indeed, CMP change mostly entails intangible learning and opportunity costs related to crop return loss risks.⁶ The presence of such learning and opportunity costs implies that, even if CMP choice can be considered as being a short run choice, switching from one CMP to another is not so straightforward. If they have no particular incentive to change – either coming from the economic context or given by public policies instruments – farmers tend to keep using the

e.g., Aubertot et al., 2007). The economic efficiency of chemical crop protection also explains the focus of plant breeding on productivity rather than on resistances during the past decades (Vanloqueren and Baret, 2009).

⁶Suitably implementing a new CMP supposes to gather information, and possible to run small scale experiments. The related costs are partly sunk when farmers finally decide to switch back from the new CMP to the usual one, for instance when implementing the CMP appears to yield unsatisfactory returns. Opportunity loss costs can occur when failing to suitably adapt the new CMP to the farm context or when the new CMP is less profitable than the usual one given wheat price levels. For instance, implementing a low-input CMP instead of a high yielding one implies lowering expected wheat yields by about 0.7 t/ha, which means an opportunity loss exceeding 140 €/ha in the yield value when wheat price exceeds 200 €/ha.

same CMP. Studying the dynamic feature of CMP choice and the determinants of CMP change is essential so that public authorities can implement an efficient policy instrument to reduce the use of chemical inputs.

Unfortunately, data sets that are usually used by agricultural production economists, which mostly come from farm accountancy data, do not contain relevant information for revealing farmers' CMP choice. The main objective of this article is to propose modelling and micro-econometric approaches for uncovering CMPs used by farmers from farm accountancy data. These data document input purchases, crop yield levels and the related prices. However, they do not provide information on important components of CMPs such as seeding dates and densities, seed variety, tillage practices, *etc.* Given that intensity in the use of chemical inputs of a CMP is directly related to the yield level targeted by this CMP, our methods aim to identify CMPs used by farmers based on their yield and chemical input use levels. Following Féménia and Letort (2016), we take advantage of the conceptual similarity between agronomists' CMPs and economists' production functions. While CMPs describe in details how various inputs and cropping operations are combined to produce a given output, production functions focus on the role of purchased inputs. Assuming that a yield production function corresponds to a given CMP, it describes how crop production levels respond to purchased input use levels when this CMP is implemented. The functional form of this function, including its parameter value in a parametric setting, accounts for the unobserved components of the considered CMP.

The pioneering work of Griliches (1957) originated an abundant literature on the adoption of agricultural production technologies, in economically developed or developing countries. Micro-econometric studies rely on data in which the adoption decisions of the considered techniques are observed. They generally focus on the adoption of specific techniques or practices (*e.g.*, use of a cultivar, tillage techniques, integrated pest management) and put emphasis on specific determinants such as learning processes and the role of uncertainty (*e.g.*, Bernard et al., 2016; Chavas and Nauges, 2020; Foster and Rosenzweig, 2010; Marra and Pannell, 2003), heterogeneity in return to adoption (*e.g.*, Michler et al., 2018; Suri, 2011) or labour constraints (*e.g.*, Fernandez-Cornejo, Hendricks, and Mishra, 2005). Specific analyses aim to assess the impacts of new technologies on yields, input uses and income (*e.g.*, Fernandez-Cornejo, 1996; Khanna, 2001; Teklewold et al., 2013).

Our use of "latent technologies" relates our modelling framework to that used in the strand of studies originated by the work of Orea and Kumbhakar (2004) and Greene (2005) on latent class stochastic frontier models (*e.g.*, Alvarez and Corral, 2010; Dakpo et al., 2021; Martinez Cillero et al., 2018). Yet, our modelling approach also intends to integrate elements that come from other modelling frameworks. First, to investigate the determinants of CMP choice we explicitly define the expected economic returns of the CMPs as key drivers of farmers' CMP choice (following here

Griliches, 1957). It casts our study into the aforementioned economic literature dealing with the adoption and the diffusion of agricultural technologies. Additionally, we assume that the current CMP choice of a farmer depends, among other factors, on its previous CMP choice. Hence, CMP choice is a dynamic process that can be modelled as a Markovian process. Our model can thus be seen as an example of hidden Markov models designed for longitudinal data (e.g., Altman, 2007; Maruotti, 2011). Moreover, we decide to account for unobserved farm heterogeneity and to consider farm specific random parameters (following the work of Koutchadé, Carpentier, and Féménia, 2018, 2020). This implies that the two main components of our modelling framework – that is to say its dynamic CMP choice sub-model and its yield and input use levels CMP specific sub-models – share common random parameters. More generally, these sub-models depend on (possibly) correlated random parameters. Accordingly, our model can be considered as a specific example of endogenous Markov switching models (e.g., Hwu, Kim, and Piger, 2019; Kim, Piger, and Startz, 2008) or of mixed hidden Markov models with heterogeneous transition probabilities (e.g., Lavielle, 2018).

To our knowledge, our study is the first combining the latent technology framework with all those elements in a single model. There are two main reasons why such modelling framework is of special interest. First, production practices are rarely documented in datasets typically used by agricultural production economists. Second, considering production practice changes is crucial when assessing the long run impacts of agri-environmental policies. Indeed, some studies tend to show that the adoption decision and its determinants are not homogeneously distributed among CMPs. For instance, Mzoughi (2011) shows that economic, social and moral concerns do not matter equally for conventional, integrated pest management (IPM) and organic farmers and Finger and El Benni (2013) find that economic incentives to adopt extensive production scheme are benefiting to already less intensive farmers. Recovering CMP might be a way (i) to better assess policy impacts that can differ across type of farms and (ii) to design more effective policy instruments according to the subpopulation we are aiming at. We illustrate our approach by investigating French farmers' CMPs for winter wheat based on a panel dataset from 1998 to 2014 covering *la Marne*, a (highly productive) arable crop production area located in eastern France. La Marne area is of special interest here because it is among the first areas in which low-input CMPs were tested with on-farm experiments (Loyce and Meynard, 1997). Moreover, la Marne is a small area characterized by rather homogeneous agro-climatic and economic conditions. We consider three main CMP types in this article: “high-yielding” CMPs (HY-CMPs), intermediate CMPs (Int-CMPs) and “low-input” CMPs (LI-CMPs). HY-CMPs can be seen as the conventional CMPs in France whereas LI-CMPs were designed as chemical inputs, especially pesticide, saving CMPs (Meynard, 1985; Meynard, 1991). Additionally, in a context of high wheat prices our simulations show that, even when imposing a 100% pesticide tax,

the share of farmer switching to low-input practices is very small. Yet, when simulating a price premium for low-input wheat producers, we observe an important increase in the share of wheat producers switching to low-input practices.

The rest of the article is organized as follows. The second section of the article discusses in more details LI-CMPs, their history and their underlying agronomic principles. From that, we derive general insights on how to account for CMP choice in economists’ production functions. The third section presents our micro-econometric framework which combines random parameters model to account for farmers’ individual heterogeneity with an hidden Markov model to answer for CMP heterogeneity. Then, we give the sketch of our estimation procedure. Fifth and sixth sections are dedicated respectively to the data presentation and to the description of the results of the exploratory analysis we performed to unveil CMPs in the data. The seventh section is dedicated to the presentation of the results from the random parameter hidden Markov model. Lastly, we discuss the obtained results and provide some concluding remarks.

3.2 Agronomic principles and brief history of “Low Input” CMPs

CMPs are defined by agronomists as ordered sequence of yield production decisions or decision rules aimed to produce a given crop. CMPs include soil preparation operations, seed variety, date and density, fertilization and pesticide applications, *etc.* We are interested in specific CMPs, the LI-CMPs proposed by agronomists in the late 1980s and, then, developed and promoted by agronomists and extension agents since the mid 1990s. LI-CMPs were developed by INRA starting in the mid 1980s and combined with multi-resistant wheat cultivars in the late 1990s (Larédo and Hocdé, 2014).

LI-CMPs can be interpreted as technological innovations aiming at providing answers to two main issues raised by conventional HY-CMPs. First, HY-CMPs are intensive in chemical input uses, which are polluting inputs. HY-CMPs are conceived to achieve high target yield levels but rely on high levels of chemical input uses, precisely because the techniques implemented for achieving high target yield levels tend to trigger the need of high fertilization and crop protection levels. Indeed, HY-CMPs aim to increase grain potential yield by increasing seeding densities, choosing early seeding dates, relying on productive seed varieties and applying large amounts of, especially nitrogen, fertilizers. Importantly, these HY techniques tend to increase pest and weed pressures and, consequently, call for efficient crop protection. Early seeding dates tend to expose crops to pest outbreaks. Nitrogen fertilizer use tends to trigger competition by weeds (Appleby, Olson, and Colbert, 1976; Henson and Jordan, 1982; Lintell Smith et al., 1992; Sexsmith and Pittman, 1963). High seed densities, productive – but susceptible to diseases – cultivars and high loads of nitrogen

fertilizer tend to increase wheat susceptibility to diseases (e.g., Boquet and Johnson, 1987; Howard, Chambers, and Logan, 1994; Roth et al., 1984). Yet, availability of efficient chemical pesticides enables farmer to control the pest and weed pressures triggered by HY techniques. The basic principle of the conception of LI-CMPs is to lower target yield levels in order to lower chemical input uses, pesticides in particular. Lowering target yield levels directly reduces crop nutrition needs and, thereby, nitrogen fertilization uses. LI-CMPs lower crop protection needs by avoiding cropping techniques that increase pest and weed pressures. Therefore, they allow reducing pesticide uses. Second, the decrease in grain prices induced by the – progressive for cereals while sudden for oilseeds – removal of the common agricultural policy (CAP) price support in 1992 called into question the profitability of grain production in the EU from the late 1990s to the mid 2000s. Due to the low grain prices during this period, HY-CMPs appeared to be much less profitable than they were in the early 2000s. The price support implemented by the CAP until the so-called McSharry reform in 1992 led most agricultural scientists to develop HY-CMPs to be adopted by European grain producers. Indeed, due to the relative scarcity of arable land in Western Europe, adopting HY-CMPs appeared to be the most profitable technological option for farmers, especially considering the price of pesticides which was rather low, to benefit from high grain prices (Mahé and Rainelli, 1987; Meynard, 1991). Despite a constant price for pesticides, lower grain prices can be an incentive to LI-CMPs.

The HY-CMPs and LI-CMPs considered by agronomists vary across time and production areas, depending on economic and agro-climatic conditions (Bouchard et al., 2008; Loyce and Meynard, 1997; Loyce et al., 2008, 2012; Rolland et al., 2003). On average, yield levels obtained with LI-CMPs are 10% lower than those obtained with HY-CMPs. Nitrogen fertilizer loads decrease by 10% from the HY-CMPs to the LI-CMPs while the use of (mostly) fungicides and insecticides is reduced by around 30%. Finally, due to the lower sowing densities in LI-CMPs seed uses decrease by around 50% when using these CMPs. Also, hardy wheat cultivars are complementary to the agronomic principles underlying the design of LI-CMPs (Larédo and Hocdé, 2014; Loyce et al., 2008). These cultivars are resistant to multiple diseases but slightly less productive than the ones typically used in HY-CMPs.

Finally, LI-CMPs are labor and fuel saving thanks to their lower expected pesticide application numbers. Hence, even if (i) no data exist on the adoption of LI-CMPs by French farmers and (ii) farm accountancy data do not contain any indicator enabling us to identify farmers using LI-CMPs, it is possible to track the economic conditions that were more or less favorable to the adoption of LI-CMPs by farmers. In particular, economic conditions tended to favor the adoption of LI-CMPs from the late 1990s to 2006, mostly due to the low grain prices observed during this period and stable prices for fungicides and insecticides from 1990 to 2016. On the contrary, the high grain price levels observed since 2007 have tended to favor conventional HY-CMPs, although

these effects of high grain prices on the profitability of HY-CMPs are partially offset by the high levels of fuel and fertilizer prices.

3.3 A Random Parameter Hidden Markov Model for modelling production choices accounting for CMPs

3.3.1 Crop production models accounting for CMP choices

The differences between the low-input and high-yielding CMPs described in the previous section suggest that these different CMPs need to be considered as different crop production technologies. A single production function – a function that mostly describes how yield levels respond to input uses – cannot account for the variety of CMP responses to input uses. In the absence of information characterizing CMPs, the CMP choice needs to be considered as a latent variable. Even if they are mostly characterized by their target yield levels and their congruent chemical input use levels, observed yield levels and chemical input use levels do not contain sufficient information for uncovering CMPs (as confirmed by the exploratory analyses). Indeed, farmers' yield and chemical input use levels are impacted by factors and events that make their direct comparison across farms largely irrelevant for identifying the CMPs that generated these levels. For instance, a farm with good soils generally obtains higher yields and uses more chemical inputs than a farm with moderate quality soils if both farms use the same CMP. A high-yielding CMP may target 9.5 tonnes per hectare of wheat in a good plot while it may only target 7.5 tonnes per hectare in a plot with poor soil quality. Similarly, climatic events as well as pest and disease infestations can impact yields and input uses in ways that significantly differ across farms. Across farm heterogeneity has significant effects on crop production and on farmers' choices even within areas of limited size (*e.g.*, Koutchadé, Carpentier, and Féménia, 2018, 2020). This justifies the choice of a random parameters model.

Yet, combining a model with latent technology and random parameters rises challenging identification issues. The fact that farmers are observed for several consecutive years (at least 3 in our dataset) plays a major role in the identification strategy of our empirical approach. Our model assumes that the observed series of input use and yield levels are generated each year by a single CMP and that farmers can change the CMP they use across time according to a Markov process. First asset of Markov processes is that they evolve relatively smoothly. Plus, modelling explicitly farmers' CMP choices sequence as Markov process allows to disentangle the effects of unobserved random events from a change in CMPs. As for the effects unobserved heterogeneity (*e.g.*, soil quality), they are disentangled from those of CMPs by assuming that production conditions are persistent at the farm level, with a fixed probability distribution at the farm population level, while CMPs can evolve over time. Next sections are dedicated to the presentation of our modelling choices

so that we can identify in our production function what can be attributed to (i) latent CMPs, (ii) unobserved heterogeneity or (iii) unobserved random events.

3.3.2 Latent CMPs models

Accounting for the specific features of CMPs and for their use in farmer production choice models requires a specific framework, even if our panel dataset is quite rich owing to its reporting cost accounting elements as well as its length and size. Let y_{it} denote the wheat yield level of farmer i in year t , and let $\mathbf{x}_{it} = (x_{j,it} : j \in J)$ denote the related vector of chemical input uses, where $J = \{1, \dots, J\}$ is the considered set of inputs. As our dataset is a unbalanced panel, we also need to define $H_i = \{t(i), \dots, T(i)\}$ as the observation period of farmer i , where $i = 1, \dots, N$ and $H_i \subseteq \{1, \dots, T\}$. We assume that farmers can produce wheat by using a CMP among the C ones collected in set $C = \{1, \dots, C\}$. CMP indices, $c \in C$, are ordered such that CMP 1 is the most intensive CMP – in the sense that it is designed to achieve the highest target yield level and, thus, relies on the highest chemical input use levels – while CMP C is the least intensive one – *i.e.* the one that relies on the lowest chemical input use levels for achieving the least target yield level. The CMP used by farmer i in year t , denoted by $r_{it} \in C$, is unobserved. Accordingly, variable r_{it} is considered as latent in our modelling framework.

Variable y_{it}^c denotes the wheat level obtained by farmer i considering that this farmer used CMP c in year t . Vector $\mathbf{x}_{it}^c = (x_{j,it}^c : j \in J)$ denotes the corresponding input use levels. The model chosen for the $(y_{it}^c, \mathbf{x}_{it}^c)$ vectors is given by

$$\begin{cases} y_{it}^c = b_{y,i}^c + d_{y,t,0} + \delta'_{y,0} \mathbf{z}_{it} + \varepsilon_{y,it}^c \\ \mathbf{x}_{it}^c = \mathbf{b}_{x,i}^c + \mathbf{d}_{x,t,0} + \Delta_{x,0} \mathbf{z}_{it} + \varepsilon_{x,it}^c \end{cases}, \text{ for } c \in C, \quad (3.1)$$

where vector \mathbf{z}_{it} contains farm characteristics (*e.g.*, arable land area, capital stock). Year specific fixed effects $\mathbf{d}_{t,0} = (d_{y,t,0}, \mathbf{d}_{x,t,0})$ capture the effects of factors or events that mostly vary across years.⁷ The estimates of year specific terms $\mathbf{d}_{t,0}$ can be used for uncovering the effects of price ratios on the CMP specific yield and input use levels. We assume that \mathbf{z}_{it} impacts, through matrix $\Delta_0 = (\delta'_{y,0}, \Delta_{x,0})$, yield and input use levels in ways that depend neither on farms and nor on CMPs. Similarly, we assume that the effects of factors or events that mostly vary across farms, which are modelled through parameters $\mathbf{d}_{t,0}$, depend neither on farms and nor on CMPs.⁸

⁷For instance, features of meteorological events or technological changes (such as the ones included in pesticides or in seed varieties) that impact the whole farm sample. These year effects also capture the effects of crop and input prices, implying that our reduced form models share common features with dual models of crop production choices that are widespread in the agricultural production literature.

⁸These homogeneity assumptions are admittedly restrictive as latent vector netput levels $\mathbf{q}_{it}^c = (y_{it}^c, \mathbf{x}_{it}^c)$ are related to production functions that may significantly differ across CMPs. This assumption is imposed mainly for practical reasons. First, identifying CMP specific effects of year specific factors is difficult in a latent CMP framework. Second, as will be seen below, this assumption significantly

Farm and CMP specific terms $\mathbf{b}_i^c = (b_{y,i}^c, \mathbf{b}_{x,i}^c)$, where $\mathbf{b}_{x,i}^c = (b_{x,j,i}^c : j \in J)$, account for the effects of CMP and production conditions on wheat yields and input uses. Term $b_{y,i}^c$ is designated, for short, as the wheat target yield level of CMP c as this practice is implemented by farmer i . Similarly, term $b_{x,j,i}^c$ is designated as the requirement in input j of CMP c . Error terms $\boldsymbol{\varepsilon}_{it}^c = (\varepsilon_{y,it}^c, \boldsymbol{\varepsilon}_{x,it}^c)$ capture the effects of random events on wheat yield and input use levels that may depend on farms, years and CMPs.

We assume that vectors \mathbf{b}_i^c , \mathbf{z}_{it} and $\boldsymbol{\varepsilon}_{it}^c$ are mutually independent.⁹ We also assume that error terms $\boldsymbol{\varepsilon}_{it}^c$ are independent across farms and years. These assumptions imply that vectors \mathbf{q}_{it} are independent across time conditionally on \mathbf{b}_i^c and \mathbf{z}_{it} . Finally we assume that $\boldsymbol{\varepsilon}_{it}^c$ is normally distributed, with $\boldsymbol{\varepsilon}_{it}^c \sim N(\mathbf{0}, \boldsymbol{\Sigma}_0^c)$.

To ensure that our latent CMP framework empirically identifies CMPs, we adopt a specific parameterization of the random parameters of our model based on the relative properties of high-yielding *versus* low-input CMPs. This parameterization defines terms $b_{y,i}^c$ and $b_{x,j,i}^c$ based on simple recursive schemes, $b_{y,i}^c = a_{y,i}^c b_{y,i}^{c-1}$ and $b_{x,j,i}^c = a_{x,j,i}^c b_{x,j,i}^{c-1}$. It implies that $b_{y,i}^c$ and $b_{x,j,i}^c$ are given by the following simple formulae:

$$b_{y,i}^c = b_{y,i}^1 \prod_{d=2}^c a_{y,i}^d \quad , \text{ for } c \in C_a,$$

and

$$b_{x,j,i}^c = b_{x,j,i}^1 \prod_{d=2}^c a_{x,j,i}^d \quad , \text{ for } c \in C_a \text{ and } j \in J,$$

where $C_a = \{2, \dots, C\}$. The conditions stating that $b_{y,i}^1 \geq 0$, $a_{y,i}^c \in [0, 1]$, $b_{x,j,i}^1 \geq 0$ and $a_{x,j,i}^c \in [0, 1]$, for $c \in C_a$ and $j \in J$, guarantee that expected yield and chemical input use levels $b_{y,i}^c$ and $b_{x,j,i}^c$ are non-negative and decrease in c . Hence, they ensure the identification of more or less intensive production technologies and fit our defining c as an index that decrease with target yield level (*i.e.*, CMP intensity in chemical input uses). These conditions can be enforced by using suitable probability distributions for random parameter vectors $\boldsymbol{\gamma}_i = (\mathbf{b}_i^1, \mathbf{a}_{y,i}, \mathbf{a}_{x,i})$, where $\mathbf{a}_{y,i} = (a_{y,i}^c : c \in C_a)$, $\mathbf{a}_{x,i} = (a_{x,j,i}^c : j \in J)$ and $\mathbf{a}_{x,i} = (a_{y,i}^c : c \in C_a)$.¹⁰ Under the considered assumptions, the probability distribution function of $\mathbf{q}_{it} = (y_{it}, \mathbf{x}_{it})$ conditional on \mathbf{z}_{it} , $\boldsymbol{\gamma}_i$ and $r_{it} = c$ is given by:

$$f(\mathbf{q}_{it} | r_{it} = c, \boldsymbol{\gamma}_i, \mathbf{z}_{it}) = \varphi(\mathbf{q}_{it} - \mathbf{b}_i^c - \mathbf{d}_{t,0} - \boldsymbol{\Delta}_0 \mathbf{z}_{it}; \boldsymbol{\Sigma}_0^c), \quad (3.2)$$

simplifies our CMP choice model. Third, and more importantly, the effects of CMPs are captured in other parts of the model of \mathbf{q}_{it}^c .

⁹Vector \mathbf{z}_{it} containing quasi-fixed input quantities, our assuming that \mathbf{z}_{it} is (strictly) exogenous with respect to $\boldsymbol{\varepsilon}_{it}^c$ and that \mathbf{b}_i^c and \mathbf{z}_{it} are independent appears reasonable (and is fairly standard).

¹⁰Interestingly, our initial intention was to consider terms $\mathbf{a}_{y,i}$ and $\mathbf{a}_{x,i}$ as fixed parameters, which is equivalent to imposing that $\mathbf{a}_{y,i} = \mathbf{a}_{y,0}$ and $\mathbf{a}_{x,i} = \mathbf{a}_{x,0}$ for $i = 1, \dots, N$. The purposes of this "fixed parameter" specification were (i) to secure the identification of the model parameters and (ii) to facilitate the comparison of the yield levels and chemical input uses across CMPs. Surprisingly enough, we couldn't estimate the model with fixed parameters $\mathbf{a}_{y,0}$ and $\mathbf{a}_{x,0}$ due to convergence issues while the estimated probability distributions of random parameters $\mathbf{a}_{y,i}$ and $\mathbf{a}_{x,i}$ display limited (although statistically significant) variability.

where $\varphi(\mathbf{a}; \mathbf{\Xi})$ is the probability distribution function of $N(\mathbf{0}, \mathbf{\Xi})$ at \mathbf{a} . The only thing we need so we can define the unconditional probability distribution of \mathbf{q}_{it} , is the probability distribution of $r_{it} = c$ as farmers' CMP choice is unobserved (and we already assume a probability distribution for γ_i).

3.3.3 A model with dynamic CMP choice

The following regime switching equation provides the link between the observed input use and yield levels on the one hand, and the set of their latent CMP specific counterparts:

$$\mathbf{q}_{it} = \sum_{c \in C} 1(r_{it} = c) \mathbf{q}_{it}^c = \sum_{c \in C} r_{it}^c \mathbf{q}_{it}^c, \quad (3.3)$$

where dummy variable r_{it}^c indicates whether farmer i chose CMP c ($r_{it}^c = 1$) or not ($r_{it}^c = 0$) in year t . We define a structural model for r_{it} in the sense that it explicitly describes how the characteristics of the latent CMP specific netput levels impact the CMP choice of expected profit maximizing farmers. Such structural model allows us to investigate how farmers choose which CMP to use.¹¹ Let $w_{y,t}$ denote wheat price paid to farmers and $w_{x,j,t}$ denotes the price paid by farmers for input j in year t . The return to chemical inputs of wheat production obtained by farm i is given by $\tilde{\pi}_{it}^c = w_{y,t} y_{it}^c - \mathbf{w}'_{x,t} \mathbf{x}_{it}^c$, when CMP c is used on this farm. But, if input purchase prices $\mathbf{w}_{x,t} = (w_{x,j,t} : j \in J)$, farm specific parameters γ_i and farm characteristics \mathbf{z}_{it} can safely be assumed to be known to farmers, most of the other terms that are part of returns $\tilde{\pi}_{it}^c$ are unknown to farmers at the beginning of the cropping season. Let π_{it}^c denote the expectation of $\tilde{\pi}_{it}^c$ by farmer i at the beginning of cropping season t . This expectation can be defined by $\pi_{it}^c = E[\tilde{\pi}_{it}^c | \omega_{it}]$ where ω_{it} denotes the information set of farmer i at the time he sows wheat to be harvested in year t . It is easily shown that:

$$\pi_{it}^c = E[w_{y,t} | \omega_{it}] b_{y,i}^c - \mathbf{w}'_{x,t} \mathbf{b}_{x,i}^c + \kappa_{it}, \quad (3.4)$$

where

$$\kappa_{it} = E[w_{y,t} | \omega_{it}] (E[d_{y,t,0} | \omega_{it}] + \delta'_{y,0} \mathbf{z}_{it}) - \mathbf{w}'_{x,t} (E[\mathbf{d}_{x,t,0} | \omega_{it}] + \mathbf{\Delta}_{x,0} \mathbf{z}_{it}).$$

Terms $E[\mathbf{d}_{x,t,0} | \omega_{it}]$ and $E[d_{y,t,0} | \omega_{it}]$ capture the effects of wheat prices and meteorological conditions on chemical input uses and wheat yield to be expected in year t . Importantly, term κ_{it} does not depend on the CMP used by the considered farmer, implying that this term is irrelevant for investigating farmers' CMP choice. We simply assume here that farmers rely on naive expectations with respect to the crop price, that is to say we assume that $E[w_{y,t} | \omega_{it}] = w_{y,t-1}$ (e.g., Koutchadé, Carpentier,

¹¹In the production frontier literature using latent class models, the probability of farmer i using CMP c in year t is defined either as a fixed probability parameter or as a probability function that depends on exogenous variables including farm characteristics or economic factors. Yet, this approach, which is focused on identifying the characteristics of latent variables \mathbf{q}_{it}^c , does not fit our objective to identify the CMP choice determinants.

and Féménia, 2018, 2020). It implies that $\pi_{it}^c = w_{y,t-1}b_{y,i}^c - \mathbf{w}'_{x,t}\mathbf{b}_{x,i}^c + \eta_{it}$.¹²

We do not expect farmers to change their CMP frequently, even if we expect the relative profitability levels of the CMPs under consideration to significantly vary across years. Because of transition costs, farmers are expected to tend to stick to the CMP they are used to. CMP choice can thus be considered as a dynamic process, in the sense that the current choice depends on the past ones. The sequence of CMP choices r_{it} is assumed to follow a (possibly) farmer specific Markov chain given the expected crop returns π_{it} . We use a first order Markov chain so we have that $P_i[r_{it} = d|r_{it-1}, \dots, r_{it(i)}, \pi_{it}] = P_i[r_{it} = d|r_{it-1}, \pi_{it}]$.¹³

To link the economic profitability of the considered CMPs and their choice by farmers, we define the transition probabilities of the CMP choice process as functions of expected returns π_{it} and of implicit CMP switch costs. Three main types of switch costs can be defined. First, expected returns π_{it} only consider chemical input costs. Yet, costs such as implementation costs of pesticide sprays or monitoring costs, argue for systematic differences in CMPs costs. These systematic differences have to be accounted for by the farmer when considering a CMP change. Second, CMP change entails chemical input uses adjustments together with adjustments in agronomical techniques, such as sowing dates and densities or seed cultivars. Finally, CMP choice can depend on farmers' attitude toward risk (e.g., Chavas and Nauges, 2020) or environmental issues (e.g., Howley et al., 2015). Such "behavioral differences" impact their willingness to pay for CMPs that are either seen as more risky or more environmental-friendly. To account for the fact that adjustment costs can vary across farmers and can also depend on the considered CMP, we incorporate farm and CMP specific random parameters in our modelling framework. Let vector α_i stack all the farm specific random parameters of the entire model and function $p(d|c, \alpha_i, \mathbf{w}_t)$ denote probability of farmer i using CMP d in year t conditionally on this farmer using CMP c in year $t - 1$, on random terms α_i and on expected price levels $\mathbf{w}_t = (w_{y,t-1}, \mathbf{w}_{x,t})$. This probability function is defined by $p(d|c, \alpha_i, \mathbf{w}_t) = P[r_{it} = d|r_{it-1} = c, \mathbf{w}_t, \alpha_i]$ and its assumed functional form is given by:

$$p(d|c, \alpha_i, \mathbf{w}_t) = \frac{\exp\left(\rho_i(\pi_{it}^d - \mu_i^{d|c})\right)}{\sum_{k \in C} \exp\left(\rho_i(\pi_{it}^k - \mu_i^{k|c})\right)}, \text{ for } (c, d) \in C \times C. \quad (3.5)$$

The functional form of transition probability function $p(d|c, \alpha_i, \mathbf{w}_t)$ is that of a (mixed)

¹²Considering adaptive anticipation schemes (e.g., Chavas and Holt, 1990) instead of the simple naive one slightly impact the quantitative estimation results but does not modify the main conclusions drawn from these results.

¹³The underlying assumption being that the last CMP choice, r_{it-1} , sums up the information content of the CMP choice history, $(r_{it-1}, \dots, r_{it(i)})$, that is relevant for modelling CMP choice r_{it} given π_{it} .

Multinomial Logit model. This discrete choice model describes the choice of a CMP from set C in year t by farmer i assuming that this farmer used CMP c in year $t - 1$.¹⁴ The underlying discrete choice model is an expected return maximization problem given by $r_{it} = \arg \max_{k \in C} \{\pi_{it}^k - \mu_i^{k|c} + (\rho_i)^{-1} e_{it}^{k|c}\}$. To go from this model on r_{it} to the probability functions displayed in Equation (3.5), we need to assume that the elements of error term vector $\mathbf{e}_{it}^{l|c} = (e_{it}^{d|c} | d \in C)$, which are known to farmer i but unobserved by the analyst, are mutually independent and follow a standard Gumbel distribution. Random parameter ρ_i is a positive scale parameter for the effects of error terms $e_{it}^{d|c}$. Hence, the larger ρ_i is, the more extended returns $\pi_{it}^k - \mu_i^{k|c}$ matter in the considered CMP choice. As specified in Equation (3.5), farmers' CMP choice depends on the relative expected returns of the CMPs, π_{it} , as well as on farmer specific switching costs $\mu_i = (\mu_i^{d|c} : (c, d) \in C^2)$.¹⁵

Equation (3.5) defines the transition probabilities of the CMP choice process. To be able to determine the probability function of the (unobserved) sequence of CMP choices for the sampled farmers, we also need to define the probability for a farmer to have chosen CMP c at his entrance year in the panel $t(i)$. Let $p_0(c | \alpha_i, \mathbf{w}_{t(i)})$ denote such probability given α_i and $\mathbf{w}_{t(i)}$. We assume that the functional form $p_0(c | \alpha_i, \mathbf{w}_{t(i)})$ is given by:

$$p_0(c | \alpha_i, \mathbf{w}_{t(i)}) = \frac{\exp\left(\sigma_i(\pi_{it(i)}^c - \eta_i^c)\right)}{\sum_{k \in C} \exp\left(\sigma_i(\pi_{it(i)}^k - \eta_i^k)\right)}, \text{ for } c \in C. \quad (3.6)$$

This functional form is chosen to account for the effects of the CMP relative profitability levels $\pi_{it(i)}$. It is inspired by the Multinomial Logit probability function associated to the expected profit maximization problem given by $\max_{k \in C} \{\pi_{it(i)}^k - \eta_i^k + (\sigma_i)^{-1} e_{it(i)}^k\}$. Term σ_i is a positive farm specific parameter scaling the effects of error terms $e_{it(i)}^k$ and terms $\eta_i = (\eta_i^c : c \in C)$ capture the effects of farmer specific costs or motives that tend to direct farmers' choice toward particular CMPs.¹⁶ When presenting the low-input CMPs, we emphasize that there are particular economic conditions that might discourage the adoption of such practices. Notably, high wheat prices tend to discourage the use of low-input practices. Thus, one could introduce a time trend in η_i^c to account for the fact that, depending on when farmer i arrives in the panel, his probability to adopt a specific CMP might changes.¹⁷

¹⁴This probability function ensures that terms $p_t(d|c; \alpha_i)$ strictly lie in the unit interval, and that terms $p(d|c, \alpha_i, \mathbf{w}_t)$ sum to 1 over $d \in C$.

¹⁵We impose the normalization constraints stating that $\mu_i^{1|c} = 0$ for $c \in C$. We choose CMP 1, the most intensive CMP, as the benchmark choice because we expect most farmers to use high yielding CMPs. Term $\mu_i^{d|c}$ denotes the switching cost incurred by farmer i when adopting CMP d while leaving CMP c relatively to (*i.e.*, minus) the switching cost incurred when adopting CMP 1. It is negative if adopting CMP d entails lower switching costs than adopting CMP 1 for farmer i .

¹⁶Condition $\eta_i^1 = 0$ is chosen as the normalization constraint for the elements of η_i .

¹⁷This point is discussed further in the Discussion/Conclusion section.

Combining Equations (3.5) and (3.6), we can finally define the probability function of the (unobserved) sequence of CMP choices for the sampled farmers. Let $\mathbf{r}_{(i)} = (r_{it} : t \in H_i)$ define the sequence of latent CMP choices of farmer i . Let function $P(\mathbf{r}_{(i)}|\boldsymbol{\alpha}_i, \mathbf{w}_{(i)})$ denote the probability function of $\mathbf{r}_{(i)}$ given $\boldsymbol{\alpha}_i$ and $\mathbf{w}_{(i)} = (\mathbf{w}_t : t \in H_i)$. Under our assumption set, computing the probability function $P_i(\mathbf{r}_{(i)}|\boldsymbol{\alpha}_i)$ of CMP choice sequence $\mathbf{r}_{(i)}$ given $\boldsymbol{\alpha}_i$ and $\mathbf{w}_{t(i)}$ yields:

$$\begin{aligned} \ln P_i(\mathbf{r}_{(i)}|\boldsymbol{\alpha}_i) &= \sum_{c \in C} r_{it(i)}^c \ln p_0(c|\boldsymbol{\alpha}_i, \mathbf{w}_{t(i)}) \\ &+ \sum_{t=t(i)+1}^{T(i)} \sum_{c \in C} \sum_{d \in C} r_{it-1}^c r_{it}^d \ln p(d|c, \boldsymbol{\alpha}_i, \mathbf{w}_t). \end{aligned} \quad (3.7)$$

Equation (3.7) corresponds to the likelihood of the considered CMP choice model for our panel dataset, with $\boldsymbol{\alpha}_i = (\gamma_i, \rho_i, \boldsymbol{\mu}_i, \sigma_i, \boldsymbol{\eta}_i)$. If we assume that error terms $\mathbf{e}_{it} = (\mathbf{e}_{it}^c | c \in C)$ and ε_{it} are independent, we can easily derive the unconditional likelihood associated to \mathbf{q}_{it} .¹⁸

3.4 Sketch of the estimation procedure

For estimation purpose, we consider a fully parametric version of our model. We assume here, for simplicity, that the probability density function of random parameter $\boldsymbol{\alpha}_i$ is multivariate normal, with $\boldsymbol{\alpha}_i \sim N(\boldsymbol{\alpha}_0, \boldsymbol{\Omega}_0)$, where variance matrix $\boldsymbol{\Omega}_0$ is left unrestricted. Since our model is fully parametric, we consider estimating its parameters, $\boldsymbol{\theta}_0 = (\mathbf{d}_0, \boldsymbol{\Delta}_0, \boldsymbol{\alpha}_0, \boldsymbol{\Omega}_0, (\boldsymbol{\Sigma}_0^c : c \in C))$, using the Maximum Likelihood (ML) approach. The contribution of farmer i to the sample likelihood function given $\boldsymbol{\alpha}_i$ at $\boldsymbol{\theta}$ is given by:

$$\begin{aligned} \ell_i(\boldsymbol{\theta}|\boldsymbol{\alpha}_i) &= \sum_{c_{it(i)} \in C} \cdots \sum_{c_{iT(i)} \in C} p_0(c_{it(i)}|\boldsymbol{\alpha}_i, \mathbf{w}_{t(i)}) \\ &\prod_{t=t(i)+1}^{T(i)} p(c_{it}|c_{it-1}, \boldsymbol{\alpha}_i, \mathbf{w}_t) \prod_{t=t(i)}^{T(i)} \varphi(\mathbf{u}_{it}^c; \boldsymbol{\Sigma}^{c_{it}}), \end{aligned} \quad (3.8)$$

where $\mathbf{u}_{it}^c = \mathbf{q}_{it} - \mathbf{b}_i^c - \mathbf{d}_t - \boldsymbol{\Delta} \mathbf{z}_{it}$. The related contribution to the sample likelihood function at $\boldsymbol{\theta}$ is given by:

$$\ell_i(\boldsymbol{\theta}) = \int \ell_i(\boldsymbol{\theta}|\boldsymbol{\zeta}) \varphi(\boldsymbol{\zeta} - \boldsymbol{\alpha}_i; \boldsymbol{\Omega}) d\boldsymbol{\zeta}. \quad (3.9)$$

¹⁸Such assumption implies that CMP choices on the one hand, and input use and yield levels on the other hand are independent conditional on $\boldsymbol{\alpha}_i$. Yet, return π_{it} is a function of γ_i , the random parameters characterizing the CMPs. Also, we allow switching cost random parameters $\boldsymbol{\mu}_i$ to be correlated with γ_i . Explicitly specifying the effects of $\boldsymbol{\alpha}_i$ in the CMP choice and crop production choice models allows us to control the endogeneity of CMP choices with respect to the crop production choices, which we assume to only depend on unobserved farm heterogeneity which is explicitly modelled through random parameters $\boldsymbol{\alpha}_i$.

The computation of likelihood terms $\ell_i(\boldsymbol{\theta})$ is particularly challenging.¹⁹ Hence the use of extensions of the Expectation-Maximization (EM) algorithm of Dempster, Laird, and Rubin (1977).²⁰

The Stochastic Approximate EM (SAEM) algorithm proposed by Delyon, Lavielle, and Moulines (1999) is a computationally efficient alternative to the Monte Carlo EM (MCEM) algorithm (Wei and Tanner, 1990), especially when the probability distributions involved in the likelihood function of the model belong to the exponential family (see, *e.g.*, Kuhn and Lavielle, 2005; Lavielle and Mbogning, 2014). It relies on a stochastic approximation approach for solving the E step. Let define vectors $\mathbf{q}_{(i)} = (\mathbf{q}_{it} : t \in H_i)$ and $\mathbf{z}_{(i)} = (\mathbf{z}_{it} : t \in H_i)$. The complete data of our model consists of (i) the observed variable vectors $\zeta_{(i)} = (\mathbf{q}_{(i)}, \mathbf{w}_{(i)}, \mathbf{z}_{(i)})$, (ii) the latent CMP choices sequence $\mathbf{r}_{(i)}$ and (iii) the random parameters vector $\boldsymbol{\alpha}_i$, for $i = 1, \dots, N$. The complete data log-likelihood function is the sample log-likelihood function of the joint model of the dependent and missing variables, $(\mathbf{q}_{(i)}, \mathbf{r}_{(i)}, \boldsymbol{\alpha}_i)$, given the exogenous variables of the model, $(\mathbf{w}_{(i)}, \mathbf{z}_{(i)})$. The complete data log-likelihood function at $\boldsymbol{\theta}$ of our model is given by:

$$\ln L^C(\boldsymbol{\theta}) = \sum_{i=1}^N \ln \ell_i^C(\boldsymbol{\theta} | \mathbf{r}_{(i)}, \boldsymbol{\alpha}_i), \quad (3.10)$$

where:

$$\ln \ell_i^C(\boldsymbol{\theta} | \mathbf{r}_{(i)}, \boldsymbol{\alpha}_i) = \left\{ \begin{array}{l} \sum_{c \in C} r_{it(i)}^c \ln p_0(c | \boldsymbol{\alpha}_i, \mathbf{w}_{t(i)}) \\ + \sum_{t=t(i)+1}^{T(i)} \sum_{d \in C} \sum_{c \in C} r_{it-1}^d r_{it}^c \ln p(c | d, \boldsymbol{\alpha}_i, \mathbf{w}_t) \\ + \sum_{t=t(i)}^{T(i)} \sum_{c \in C} r_{it}^c \ln \varphi(\mathbf{q}_{it} - \mathbf{b}_i^c - \mathbf{d}_t - \Delta \mathbf{z}_{it}; \boldsymbol{\Sigma}^c) \\ + \ln \varphi(\boldsymbol{\alpha}_i - \boldsymbol{\alpha}; \boldsymbol{\Omega}) \end{array} \right\}.$$

At iteration n of an EM type algorithm, the objective of the E step is to integrate $\ln L^C(\boldsymbol{\theta})$ over the probability distribution of the missing data $(\mathbf{r}_{(i)}, \boldsymbol{\alpha}_i)$ conditional on the observed data $\zeta_{(i)}$ evaluated at $\boldsymbol{\theta}^{(n)}$, the last available estimate of $\boldsymbol{\theta}_0$. This consists of computing the conditional expectations $E^{(n)}[\ln \ell_i^C(\boldsymbol{\theta} | \mathbf{r}_{(i)}, \boldsymbol{\alpha}_i) | \zeta_{(i)}]$, for $i = 1, \dots, N$.

The computation of conditional expectations $E^{(n)}[\ln \ell_i^C(\boldsymbol{\theta} | \mathbf{r}_{(i)}, \boldsymbol{\alpha}_i) | \zeta_{(i)}, \boldsymbol{\alpha}_i]$ consists of computing the conditional expectations of terms $r_{it(i)}^c$, r_{it}^c and $r_{it-1}^d r_{it}^c$. Under our model assumptions we can show that $E^{(n)}[r_{it(i)}^c | \zeta_{(i)}, \boldsymbol{\alpha}_i] = p_0(c | \boldsymbol{\alpha}_i, \mathbf{w}_{t(i)})$. Likewise, expectation terms $p_t(c | \boldsymbol{\alpha}_i, \mathbf{w}_{(i)}) = E^{(n)}[r_{it}^c | \zeta_{(i)}, \boldsymbol{\alpha}_i]$ and $s_t(c, d | \boldsymbol{\alpha}_i, \mathbf{w}_{(i)}) = E^{(n)}[r_{it-1}^d r_{it}^c | \zeta_{(i)}, \boldsymbol{\alpha}_i]$

¹⁹This problem combines two issues. As it stands in Equation (3.8) the expression of $\ell_i(\boldsymbol{\theta} | \boldsymbol{\alpha}_i)$ is of little or no computational use. It quickly becomes intractable as C or/and T grows to moderate levels. Second, the integration problem involved in Equation (3.9) can rarely be solved either analytically or numerically. Computing $\ell_i(\boldsymbol{\theta})$ requires simulation methods. Solving these issues leads to an awkward simulated sample log-likelihood function that is particularly challenging to maximize in $\boldsymbol{\theta}$.

²⁰EM type algorithms are particularly well suited for maximizing the log-likelihood functions of models involving missing variables. The latent CMP choices r_{it} and the random parameters $\boldsymbol{\alpha}_i$ of our model are examples of missing variables. This algorithm allows solving a complicated ML problem by iteratively solving a sequence of much simpler problems. Each iteration of an EM type algorithm involves an expectation (E) step, which consists of integrating a conditional expectation, and a maximization (M) step.

can be defined as functions of initial probability functions $p_0(c|\boldsymbol{\alpha}_i, \mathbf{w}_{t(i)})$ and of transition probability functions $p(c|d, \boldsymbol{\alpha}_i, \mathbf{w}_t)$. As neither functions $p_0(c|\boldsymbol{\alpha}_i, \mathbf{w}_{t(i)})$ nor functions $p(c|d, \boldsymbol{\alpha}_i, \mathbf{w}_t)$ depend on elements of fixed parameter $\boldsymbol{\theta}$, the same observation holds for functions $p_t(c|\boldsymbol{\alpha}_i, \mathbf{w}_{(i)})$ and $s_t(c, d|\boldsymbol{\alpha}_i, \mathbf{w}_{(i)})$.

Given these results and observations, computing $E^{(n)}[\ln \ell_i^C(\boldsymbol{\theta}|\mathbf{r}_{(i)}, \boldsymbol{\alpha}_i)|\zeta_{(i)}]$ consists of computing:

$$\begin{aligned} & E^{(n)}[\ln \ell_i^C(\boldsymbol{\theta}|\mathbf{r}_{(i)}, \boldsymbol{\alpha}_i)|\zeta_{(i)}] \\ &= \\ & \left\{ \begin{aligned} & E^{(n)}[R(\boldsymbol{\alpha}_i, \mathbf{w}_{(i)})|\zeta_{(i)}] \\ & + \sum_{t=t(i)}^{T(i)} \sum_{c \in C} E^{(n)}[s_t(c, d|\boldsymbol{\alpha}_i, \mathbf{w}_{(i)}) \ln \varphi(\mathbf{q}_{it} - \mathbf{b}_i^c - \mathbf{d}_t - \Delta \mathbf{z}_{it}; \boldsymbol{\Sigma}^c)|\zeta_{(i)}] \\ & + E^{(n)}[\ln \varphi(\boldsymbol{\alpha}_i - \boldsymbol{\alpha}; \boldsymbol{\Omega})|\zeta_{(i)}] \end{aligned} \right\}, \end{aligned} \quad (3.11)$$

where term:

$$R(\boldsymbol{\alpha}_i, \mathbf{w}_{(i)}) = \left\{ \begin{aligned} & \sum_{c \in C} p_0(c|\boldsymbol{\alpha}_i, \mathbf{w}_{t(i)}) \ln p_0(c|\boldsymbol{\alpha}_i, \mathbf{w}_{t(i)}) \\ & + \sum_{t=t(i)+1}^{T(i)} \sum_{d \in C} \sum_{c \in C} s_t(c, d|\boldsymbol{\alpha}_i, \mathbf{w}_{(i)}) \ln p(c|d, \boldsymbol{\alpha}_i, \mathbf{w}_t) \end{aligned} \right\},$$

does not involve any element of $\boldsymbol{\theta}$.

The expectations conditional on the observed data $\zeta_{(i)}$ involved in Equation (3.11) can be integrated using simulation methods. Whatever the simulation method, these expectations are approximated by weighted means of functions of simulations of random parameters $\boldsymbol{\alpha}_i$. Let $\hat{\boldsymbol{\alpha}}_i^{(n),j}$ denote the considered random draws of $\boldsymbol{\alpha}_i$ and $\hat{\omega}_i^{(n),j}$ their related weights, for $j = 1, \dots, J^{(n)}$, where $J^{(n)}$ denote the draw number considered for iteration n . The conditional expectation of $\ln L^C(\boldsymbol{\theta})$ is approximated by:

$$\begin{aligned} & E^{(n)}[\ln L^C(\boldsymbol{\theta})|\zeta] \\ & \simeq \\ & \sum_{i=1}^N \left\{ \begin{aligned} & E^{(n)}[R(\boldsymbol{\alpha}_i, \mathbf{w}_{(i)})|\zeta_{(i)}] \\ & + \sum_{t=t(i)}^{T(i)} \sum_{c \in C} \sum_{j=1}^{J^{(n)}} \hat{\omega}_i^{(n),j} s_t(c, d|\hat{\boldsymbol{\alpha}}_i^{(n),j}, \mathbf{w}_{(i)}) \ln \varphi(\mathbf{q}_{it} - \hat{\mathbf{b}}_i^{c,(n),j} - \mathbf{d}_t - \Delta \mathbf{z}_{it}; \boldsymbol{\Sigma}^c) \\ & + \sum_{j=1}^{J^{(n)}} \hat{\omega}_i^{(n),j} \ln \varphi(\hat{\boldsymbol{\alpha}}_i^{(n),j} - \boldsymbol{\alpha}; \boldsymbol{\Omega}) \end{aligned} \right\}. \end{aligned} \quad (3.12)$$

In our empirical application, we used the SAEM version of this E step and an importance sampling approach for integrating terms $E^{(n)}[\ln \ell_i^C(\boldsymbol{\theta}|\mathbf{r}_{(i)}, \boldsymbol{\alpha}_i)|\zeta_{(i)}]$. We used the probability density of $\mathcal{N}(\boldsymbol{\alpha}^{(n+1)}, \boldsymbol{\Omega}^{(n+1)})$ as the proposal density for the random draws of $\boldsymbol{\alpha}_i$.²¹

Solving the M step at iteration n then consists of maximizing $E^{(n)}[\ln L^C(\boldsymbol{\theta})|\zeta]$ in $\boldsymbol{\theta}$ for obtaining $\boldsymbol{\theta}^{(n+1)}$. In our case, this maximization problem can be solved in two

²¹The whole SAEM procedure and its explicit forms can be found in Appendix 3.9.1.

steps. Solving problem:

$$\max_{(\mathbf{d}, \Delta, \Sigma)} \sum_{i=1}^N \sum_{t=t(i)}^{T(i)} \sum_{c \in C} \sum_{j=1}^{J(i)} \hat{\omega}_i^{(n),j} s_t(c, d | \hat{\alpha}_i^{(n),j}, \mathbf{w}_{(i)}) \ln \varphi(\mathbf{q}_{it} - \hat{\mathbf{b}}_i^{c,(n),j} - \mathbf{d}_t - \Delta \mathbf{z}_{it}; \Sigma^c), \quad (3.13)$$

yields $(\mathbf{d}^{(n+1)}, \Delta^{(n+1)}, (\Sigma^{c,(n+1)} : c \in C))$, the first part of $\theta^{(n+1)}$.

On the other hand, solving problem:

$$\max_{(\alpha, \Omega)} \sum_{i=1}^N \sum_{j=1}^{J(i)} \hat{\omega}_i^{(n),j} \ln \varphi(\hat{\alpha}_i^{(n),j} - \alpha; \Omega), \quad (3.14)$$

yields $(\alpha^{(n+1)}, \Omega^{(n+1)})$, the second part of $\theta^{(n+1)}$. Both problems are equivalent to weighted ML problems of linear multivariate Gaussian models.

The E and M steps described above are iterated until numerical convergence (see Appendix 3.9.1 for more details on the estimation procedure).

3.5 Data

We use an unbalanced panel data set that considers, from 1998 to 2014, input uses and yields of winter wheat for a sample of farmers located in *La Marne*, a French department. These data have been extracted from cost accounting data provided by the CDER, the main accounting agency dealing with farming operations in the considered area. Among this cost accounting data are the received wheat prices. As our approach requires to build price anticipation to evaluate the anticipated revenue associated to each technology, data from year 1998 was dropped from the final data set so we can build the anticipated wheat price variable. Additionnally, farms that were observed less than four times in the panel were dropped. This constraint comes from the fact we are using a model with random parameters, *i.e.* we need to observe each farms multiple times so we can estimate those random parameters. Overall, the data set gathers 1351 farmers that are observed for 10 years on average. Number of farms observed each year is reported in Table 3.1. Number of observed farms is steady between 2002 and 2009 and is more fluctuating at the beginning and at the end of the time period. To evaluate the effect of the yearly changing composition of the panel, the table also provides the average size of the farms and of the winter wheat cropped surface. Farm size and cropped wheat surface tend to increase over years. The phenomenon of increased farm size is not specific to our data, it is a global trend in France that is highlighted by Agreste (see Poullette, 2018). Yet, the share of the surface allocated to winter wheat is stable, around 30%. Provided that farm size is expected to affect technology choice through time constraint, the normalized arable land area was integrated in the set of \mathbf{d}_{it} .

TABLE 3.1: Annual mean (and standard deviation) of arable land area and wheat acreage, from 1999 to 2014

Year	N	Arable land area (in ha)	Wheat acreage (in ha)
1999	678	172.61 (86.54)	52.40 (29.31)
2000	824	179.11 (89.92)	57.42 (33.41)
2001	829	176.60 (90.98)	53.91 (30.33)
2002	959	181.89 (92.45)	56.42 (33.24)
2003	946	184.96 (92.36)	53.74 (32.49)
2004	924	185.40 (93.46)	55.10 (33.32)
2005	923	184.84 (91.20)	57.12 (33.72)
2006	943	188.66 (97.36)	58.29 (35.69)
2007	959	191.10 (99.18)	59.06 (36.38)
2008	933	193.10 (101.04)	60.52 (38.09)
2009	930	193.67 (99.40)	58.52 (36.99)
2010	823	191.34 (95.56)	60.51 (34.40)
2011	651	186.83 (87.57)	61.86 (36.12)
2012	773	196.33 (97.97)	55.85 (36.81)
2013	739	198.32 (103.19)	65.40 (38.45)
2014	694	202.06 (102.77)	64.19 (39.08)

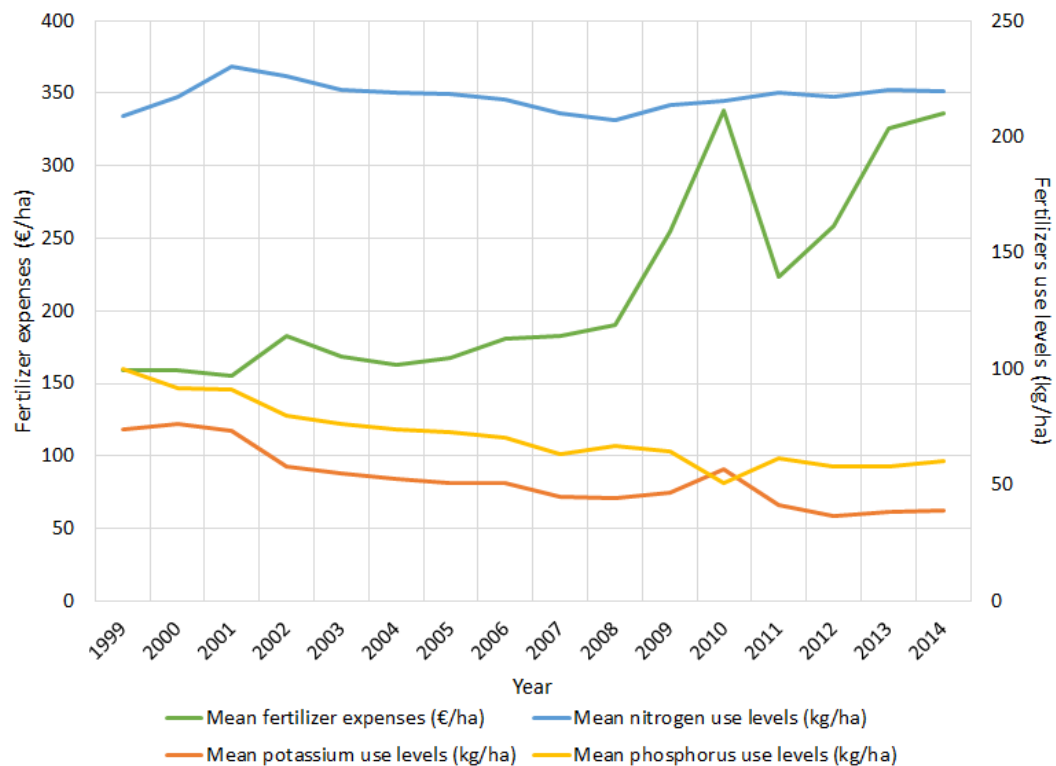
Source: CDER data.

Apart from received wheat prices, our application primarily makes use of recorded winter wheat yields and fertilizer and pesticide expenditures devoted to wheat production. Because pesticide uses are not available in the data (and rarely are at the farm level), pesticide expenditures are used as proxy for pesticide uses, even if recent studies tend to show that the correlation between input expenditures and input use is relatively low (*e.g.*, Uthes et al., 2019). Even if fertilizer uses are recorded in the data, to treat both inputs symmetrically we also use a fertilizer expenditure variable. Yet, as the quantities of the Big-Three nutrients – nitrogen (N), phosphorus (P) and potassium (K) – are available in the dataset, the correlation between expenditures and uses can be investigated.

Figure 3.1 shows that until 2008-2009, fertilizer expenditures and NPK quantities seem to be well correlated. Then, in 2008, we observe a decrease in fertilizers expenses before they peak in 2009. On the other hand, quantities of NPK are quite steady over the whole period. Such disconnection between fertilizer uses and expenditures can be explained by the 2007 economic crisis. During that period, oil prices have increased and have impacted fertilizer prices (*e.g.*, Baffes, 2007; Chen et al., 2012) and farmers might have delayed their fertilizer expenditures in 2008, relying on their stock, and have had to compensate in 2009. To overcome such pitfall, NPK quantities were regressed on fertilizers expenses to get price proxies that account for inventories effect. The build price index corresponds to the nutrient price associated to the global fertilizer expenses given the quantities of nutrients that were really used by the farmer. The nitrogen fertilizer expenditures is better related to the used

quantities than the original fertilizer expenditure variable and thus will be used in the main analysis as proxy for fertilizer uses.

FIGURE 3.1: Annual mean fertilizers expenses and nutrients uses from 1999 to 2014

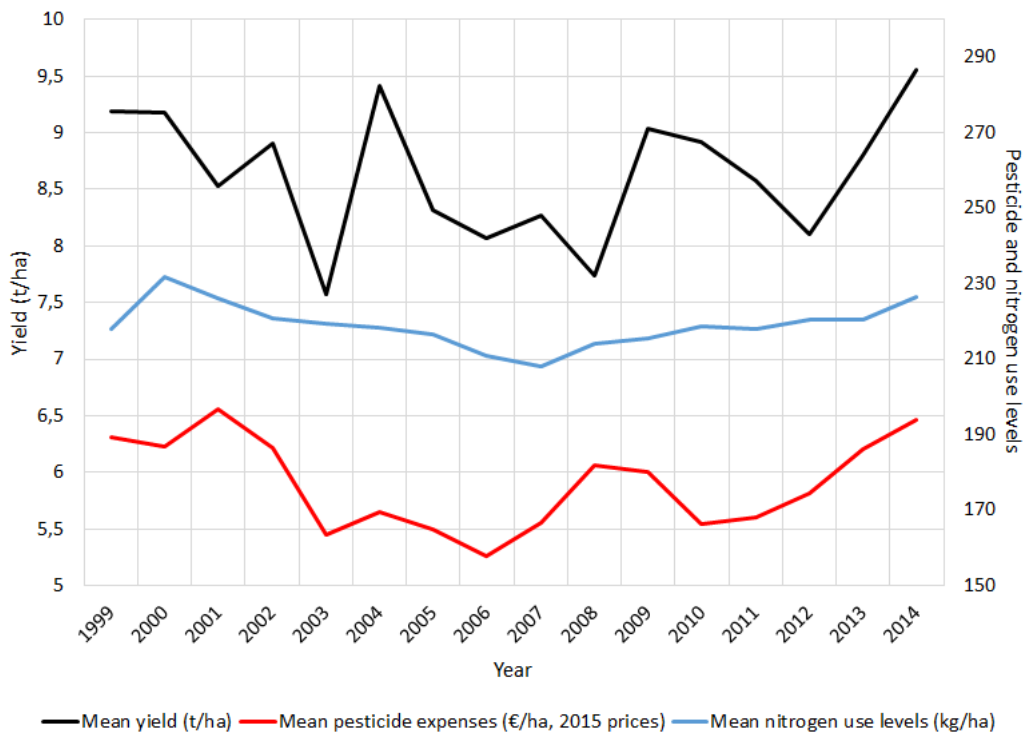


Source: CDER data.

Mean yield, fertilizer and pesticide expenses are reported in Figure 3.2. Fertilizer expenses are steadier than pesticide expenses. This might be due to the fact that pesticide uses are dependent from weather conditions and pest invasion, *i.e.* they are more likely to vary across time. Yet, the most time dependent variable remains yields that are deeply subjected to weather conditions. They vary from 7.57 tonnes per hectare in 2003 to 9.55 tonnes per hectare in 2014 with an overall mean around 8.6 tonnes per hectare.

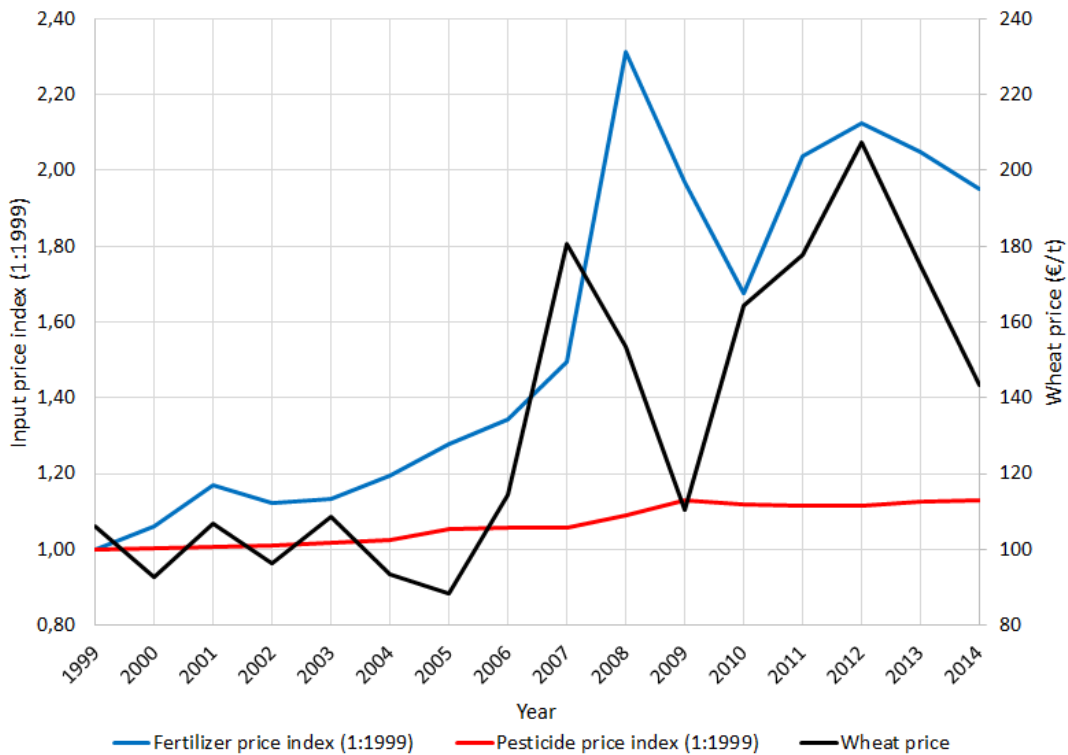
As for pesticide price indices, they were obtained from the French department of Agriculture. Figure 3.3 represents the different price variables. Pesticide prices are quite steady whereas wheat and nitrogen prices are following the same trend. They increase after 2007 crisis and are volatile since then. Given the economic context prevailing from 1999 to 2007, *i.e.* low wheat prices, and the promotion process of LI-CMPs during this period, we expect a small share of farmers using LI-CMPs in the late 1990s and an increase in this share until 2006.

FIGURE 3.2: Annual mean yields and input expenses from 1999-2014



Source: CDER data.

FIGURE 3.3: Wheat, pesticide and nitrogen annual prices from 1999-2014



Sources: IPAMPA data from Agreste/INSEE for pesticides prices ; calculations on CDER data to get mean wheat prices and a proxy for nitrogen prices.

The effects on the adoption of LI-CMPs of the combination of high wheat and fertilizer prices since 2007 is difficult to predict, notwithstanding the effects of these prices volatility. If the wheat price effects dominate those of the input prices, it is possible that farmers having adopted LI-CMPs switch back to HY-CMPs. Results from our HMM framework might help us weight the effect of such economic considerations on the adoption of low-input practices.

3.6 Insights from “exploratory” analyses

3.6.1 Presentation of the exploratory approaches

Before presenting the results of the hidden Markov model, we provide some insights based on exploratory analyses that were performed on the data. We called them “exploratory” because these analyses are simpler and are performed on sub-samples. Their purpose is to confirm the coexistence of CMPs in the data. If the interest of differentiating CMP choice has been evoked previously, nothing guarantees that different CMPs do coexist among our French farmers of the Marne area.

Clustering analysis, generally defined as the “art of finding groups in data” (Kaufman and Rousseeuw, 2009), might allow us to distinguish for farmers’ CMPs based on their yield and input use levels. If we define by $\mathbf{s}_i = (y_i, \mathbf{x}_i)$ the vector of yield and input uses from farmer i , we need to define a distance metric d to evaluate how similar (respectively dissimilar) are farmers i and j based on their characteristics \mathbf{s}_i and \mathbf{s}_j . Clustering algorithms aim at maximizing the dissimilarity between the built groups. Suppose that there are $c = 1, \dots, C$ groups among the data, each group being characterized by a vector of average yield and input uses $\boldsymbol{\mu}_c$. Then, farmer i is allocated to its closest group c , *i.e.* $c_i = \arg \min_{c \in C} d(\mathbf{s}_i, \boldsymbol{\mu}_c)$. This is the general principle of clustering algorithms. Depending on the type of algorithm we use – hierarchical and partitioning are the two main algorithm types – the stop criterion is different (see Appendix 3.9.2 for a description of the k-means and AHC algorithms).

Standard clustering models are static, *i.e.* there is no cluster switching. This is not an issue when considering cross-sectional data but can become one when having panel data as we do. In the panel data case, we have two main possibilities to perform a clustering analysis. First, we can perform the clustering analysis by integrating multiple time periods in the algorithm. Thus, we assume that for the length of the time period considered, an individual belongs to the same group. Another possibility is to consider the yearly data and perform a clustering analysis for each year separately. Yet, by doing so, the group belonging is independent between years. Both situations generate pitfalls in the case of technology adoption. Indeed, the larger the period considered, the less the assumption that a farmer has the same technology implemented on the whole period might stand. On the other hand, assuming that technology adoption is independent at each period is unsatisfactory as well. CMP

choice can be considered as a rather short term choice that does not require heavy investment in terms of machinery. However, it is highly unlikely that CMP choice observed at time t is not correlated to the CMP choice done at time $t - 1$. Moreover, by performing cluster analysis on yearly yields and input levels, the risk would be to separate farmers only based on their efficiency.

We choose an in-between solution which is to divide the data set into 4 years sub-panels.²² The 4-year length was chosen arbitrarily but seems a rather good compromise between (i) a too short period not permitting to identify underlying structure in the conjoint evolution of yield and input use levels, (ii) a too long period that will endanger the hypothesis of CMP stability. Additionally, to avoid the problem of missing data, we only consider farmers that were observed during the four considered years. As input uses - pesticides and fertilizers - and yield levels were observed multiple times - once each year - a principal component analysis was run to get summarized information and denoise data (Husson, Josse, and Pagès, 2010). Clustering analysis was then performed on the resulting principal components.

The second step of the exploratory approach consists in the use of a latent class model (see Bartholomew, Knott, and Moustaki, 2011). Latent class modelling consists in specifying and estimating statistical models of farmers’ yield supply and input demand functions while allowing farmers to use CMPs to be selected from a finite set of “latent CMPs”. Let define a simple yield supply and input demand functions that depend on the implemented CMP c :

$$\begin{cases} y_{it} &= b_{y,t}^c + \varepsilon_{y,it}^c \\ \mathbf{x}_{it} &= \mathbf{b}_{x,t}^c + \boldsymbol{\varepsilon}_{x,it}^c \end{cases} ,$$

with $b_{y,t}^c$ and $\mathbf{b}_{x,t}^c$ representing respectively the objective yield and input uses associated to CMP c given the weather conditions observed during year t . $\varepsilon_{y,it}^c$ and $\boldsymbol{\varepsilon}_{x,it}^c$ correspond to the stochastic part of yield and input uses that are creating a gap between the targeted levels and the observed ones. Assuming that (i) $\boldsymbol{\varepsilon}_{it}^c = (\varepsilon_{y,it}^c, \boldsymbol{\varepsilon}_{x,it}^c)$ and $\mathbf{b}_t^c = (b_{y,t}^c, \mathbf{b}_{x,t}^c)$ are independent and (ii) the probability density function of $\boldsymbol{\varepsilon}_{it}^c$ is known up to a parameter to be estimated, λ_0^c , the probability density function of $\mathbf{s}_{it} = (y_{it}^c, \mathbf{x}_{it}^c)$ conditionally on \mathbf{b}_t^c and on farmer i choosing CMP c can be derived. Let $\nu(\mathbf{s}_{it}, \mathbf{b}_t^c; \lambda_0^c)$ denote this probability density function. This model is quite standard as long as the CMP c is observed. When technology is latent, as in our case, writing the probability density function $\nu(\mathbf{s}_{it}, \mathbf{b}_t^c; \lambda_0^c)$ requires to recover the distribution of the

²²Sub-panels were built on a restrictive data set. We considered only farms that were located in *La Marne crayeuse area*, which is characterized by homogeneous agro-climatic and economic conditions. We choose to restrict ourselves to this area because effects of exogeneous factors such soil fertility and climate on yields and chemical inputs can be confounded with CMP choices, especially as we do not give any structure to those CMP choices. Restricting ourselves to this area for the exploratory analyses prevents us for such confusion.

latent variable c_i . Here, to keep things simple in this exploratory approach, we look for $\pi_i^c = P[c_i = c]$ and do not introduce covariates in the discrete choice of c_i .²³ The modelling assumptions described above yield that the probability density function of sequence $s_{i(h)}$ conditionally on $\mathbf{z}_{s,i}$ (*i.e.*, ignoring the unobserved CMP choice of farmer i) is given by:

$$\mu(\mathbf{s}_i, \mathbf{z}_{s,i}; \lambda_0) = \sum_{c \in \mathcal{C}} \pi_i^c v(\mathbf{s}_i, \mathbf{z}_{s,i}; \lambda_0^c).$$

This probability density function is a suitable basis for estimating parameters based on a maximum likelihood approach. In particular, we use the Expectation-Maximization algorithm (Dempster, Laird, and Rubin, 1977). As for the clustering analysis, this simplified latent class model was performed on the 4 years sub-panels and we hold the assumption that the CMP class is constant over the period.²⁴

3.6.2 Insights from results and comparison with the HMM approach

In this subsection, we present the results from these exploratory approaches and the insights they provide on CMP choice. First, we decide to check for the coexistence of a low-input and a high-yielding CMP among the French wheat producers. The results we show in there are results from the latent class model.²⁵

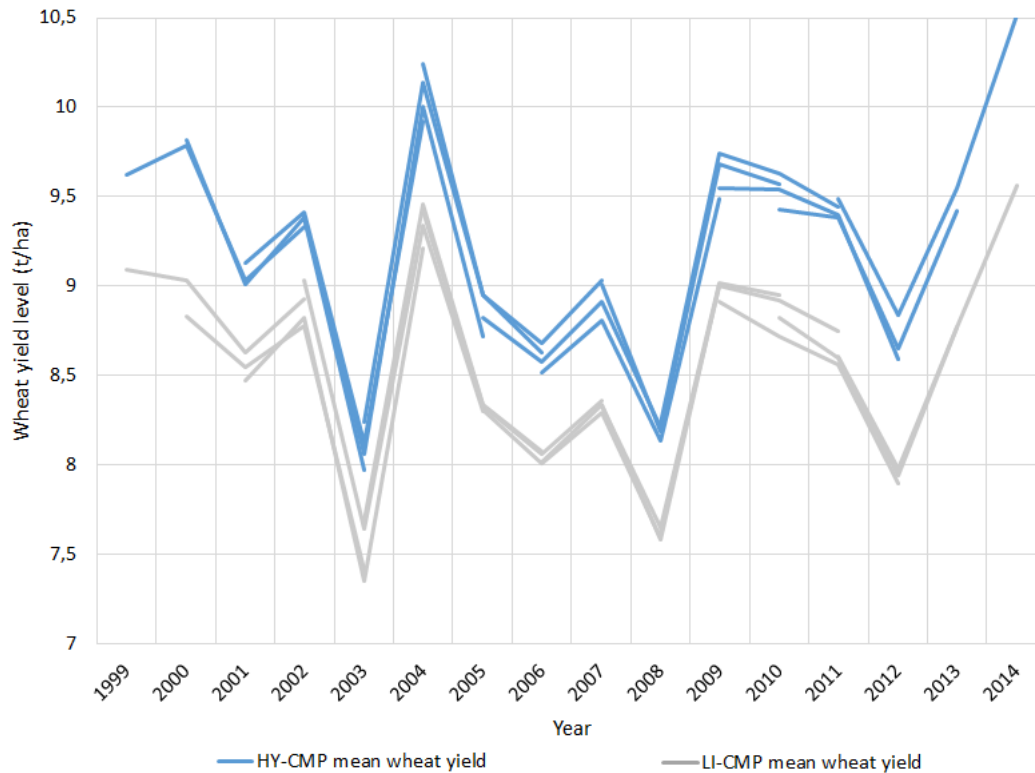
Several points can be drawn from Figures 3.4 and 3.5. The first is the relative stability of the built groups across sub-panels. As the latent class model was run independently on each sub-panels, we could have fear that their composition across sub-panels vary a lot. The fact that the groups' characteristics in terms of yield and input expenses are not "bouncing" between sub-panels is reassuring and might indicate that we have distinguished for different production practices. The second interesting fact is the parallel trend we observe through time between the two groups. Indeed, either for yield or input use levels, low-input and high-yielding groups seem to react in the same way to weather and economic conditions. Such observation strengthens the modelling choice we have made for the HMM relatively to the year specific fixed effect that are common to all CMPs, assuming that the groups we observed here can be considered as CMP classes.

²³The estimation procedure will remain the same when introducing covariates. The only difference would be that we need to introduce a parametric model for c_i and write the probabilistic discrete choice model of c_i as $r_c(\mathbf{z}_{r,i}; \mu_0) = P[c_i = c | \mathbf{z}_{r,i}]$ with $\mathbf{z}_{r,i}$ the covariates that are affecting the choice of c_i . In our case, it will correspond to CMP choice potential drivers.

²⁴Given the basic specification we use for \mathbf{s}_i and because we do not include covariates in c_i choice, this latent class model can be compared to a partitioning approach with the likelihood as a distance metric. The major difference is that we suppose here a multivariate Gaussian distribution whereas in our clustering analysis we apply no formal constraint on the data.

²⁵The EM algorithm requires initial values for parameters (\mathbf{b}_i^c, π^c). We use the clustering result from an AHC analysis to define them. We assess the results robustness to this set of initial values by using the results from a k-means analysis. Results can be found in Appendix 3.9.3.

FIGURE 3.4: Estimated annual mean yield of high-yielding and low-input farmers, from 1999 to 2014



Source: Authors' calculations on CDER data.

Results in terms of estimated yield and input use levels for the two clusters are also very interesting and in line with the CMPs characteristics as described by agronomists. On average, we observe that LI-CMPs yield level is lower by 7% in comparison to the HY-CMPs one. As for input expenses, differences are of 5% for nitrogen and 20% for pesticide expenses between LI-CMPs and HY-CMPs. Such difference in yields combined to the difference in input uses tends to show that our groups do not distinguish for farmer efficiency but for different production practices. The assumption that observed groups correspond to CMP groups is reinforced by the fact that the average differences we found between what appear to be a low-input groups and a high-yielding one are coherent with the figures given by Loyce and Meynard (Loyce and Meynard, 1997; Loyce et al., 2008, 2012) when experimenting low-input practices.

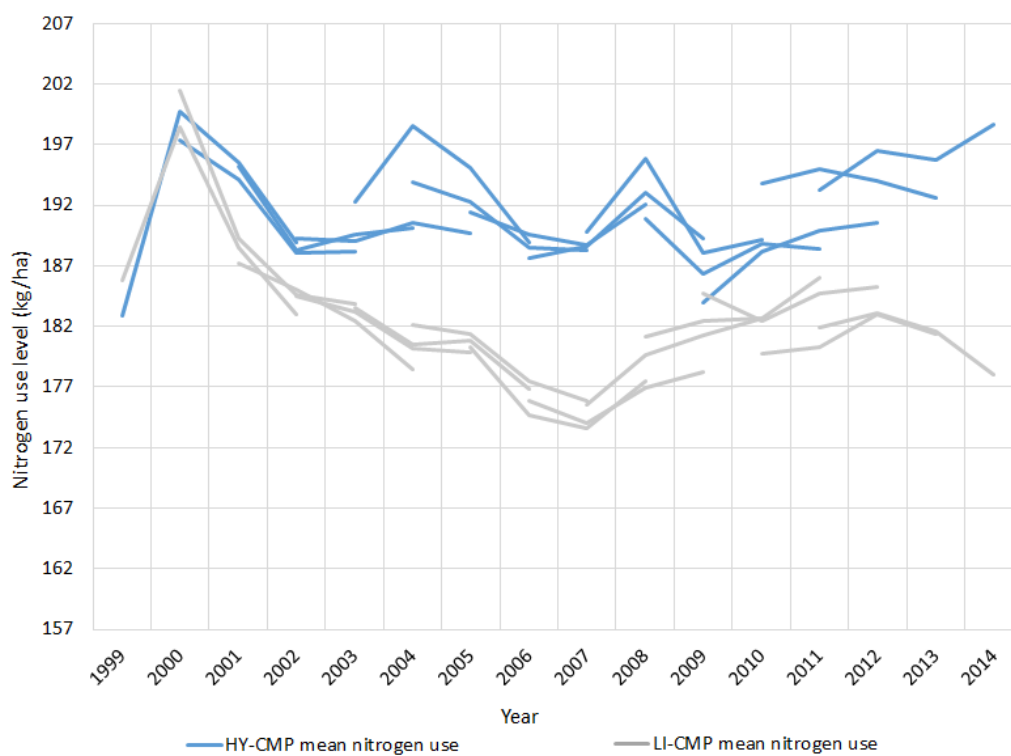
FIGURE 3.5: Estimated annual mean input uses of high-yielding and low-input farmers, from 1999 to 2014

(A) Estimated annual mean pesticide expenses in high-yielding and low-input CMPs



Source: Authors' calculations on CDER data.

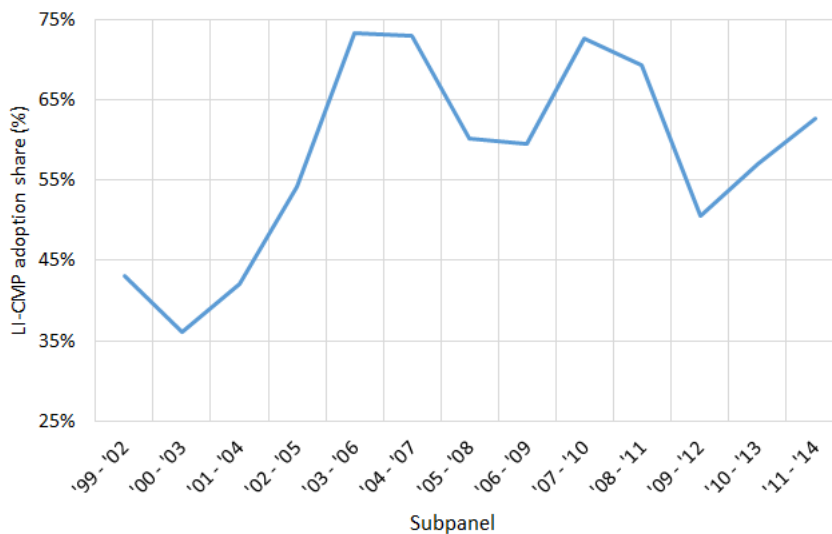
(B) Estimated annual mean nitrogen uses in high-yielding and low-input CMPs



Source: Authors' calculations on CDER data.

Regarding the size of the LI-CMP cluster over time, results are very encouraging as well. LI-CMP adoption is expected to be favored by economic conditions in the late 1990s until 2007 approximately. However, with the 2007 crisis and the associated increase in wheat prices, attractiveness of low-input practices should decrease by then. This is the global trend we can observe in Figure 3.6, with a slight increase of low-input attractiveness at the end of the period.²⁶

FIGURE 3.6: Estimated share of farmers who adopted a low-input CMP, from 1999 to 2014



Source: Authors' calculations on CDER data.

Overall, the consistency of the results obtained with the “exploratory” approaches across time and methodology (see Appendix 3.9.3 for clustering results) plus the conformity of the LI and HY clusters' characteristics to the literature are supportive elements for the use of latent class methods on the data set. Yet, the size of the low-input group we obtained here is quite surprising as they represent more than 50% of farmers for most time periods. Even if such practices were experimented during those years in this area by INRA, we doubt that low-input practices were adopted by up-to three-quarter of farmers. The hypothesis we can formulate here is that the latent class model allows us to distinguish farmers with very high-yield targets from farmers who are targeting slightly lower yields and “true” low-input farmers.

Thus, we tried to compute the same latent class model with three classes, as to distinguish for farmers who have an “intermediate” CMP and the real low-input

²⁶This result tends to strengthen the idea that, in our HMM framework, we should consider including a time trend in the initial probability to adopt a CMP c for a farmer. This choice is justified by the previous economic considerations on CMP adoption and is reinforced by the results of the exploratory analyses which tend to show that adoption of low-input practices vary with time.

farmers. Yet, the results we obtained with the latent class model for three classes are disappointing. Figure 3.7 show that, as for yield levels, results are quite time consistent.

FIGURE 3.7: Estimated annual mean yields of high-yielding, intermediate and low-input CMPs, from 1999 to 2014

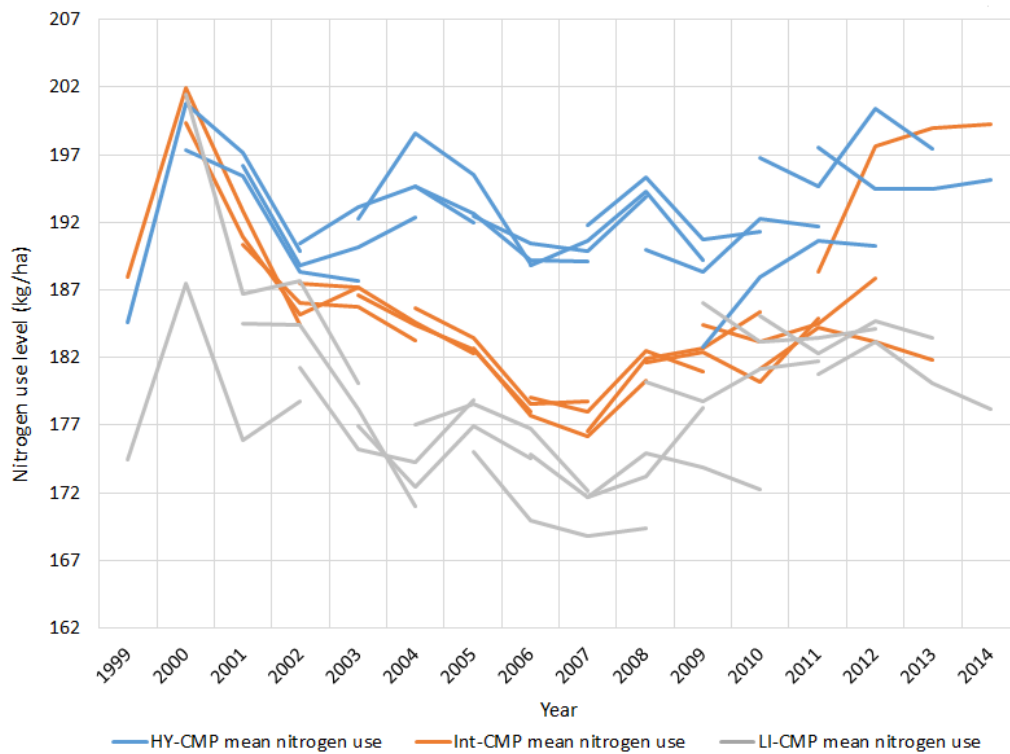


Source: Authors' calculations on CDER data.

Yet, Figure 3.8 displays a lack of consistency when considering fertilizer and pesticide expenses. In particular, we can observe some “bounces” between sub-panels. Additionally, if differences in yield are effective, for input expenses we can observe some group collapsing especially for the last sub-panels. Plus, what we called the LI-CMP has greater nitrogen expenses than the intermediate CMP for lower yield levels. It might indicate that we distinguish here for farmers' efficiency rather than production practices. Hence, the results from this three-class approach cannot be considered as uncovering CMPs' groups. The “exploratory approaches” display their limits and justify the use of a more structured model as the hidden Markov model we presented previously. Especially, we put constraints to make sure we distinguish for CMPs rather than for efficiency groups or other heterogeneity factors.

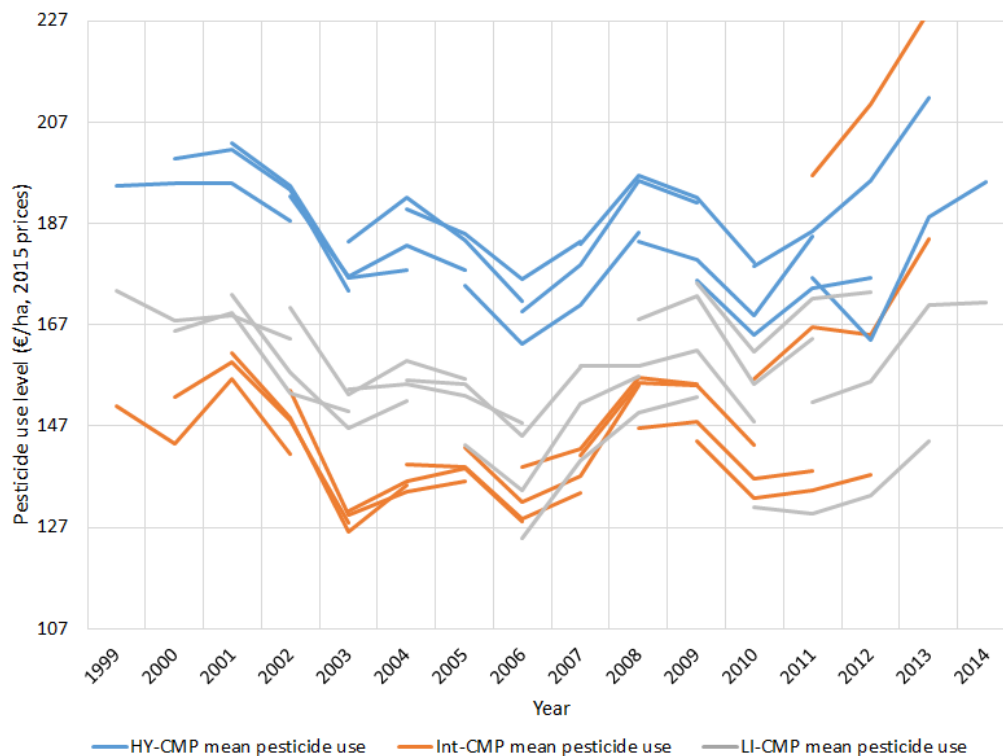
FIGURE 3.8: Estimated annual mean input uses of high-yielding, intermediate and low-input CMPs, from 1999 to 2014

(A) Mean nitrogen expenses in high-yielding, intermediate and low-input CMPs



Source: Authors' calculations on CDER data.

(B) Mean pesticide expenses in high-yielding, intermediate and low-input CMPs



Source: Authors' calculations on CDER data.

3.7 Results

In this section we present the results from the random parameter hidden Markov model (RPHMM). Before presenting the characteristics of the resulting CMP categories, we should check for the obtained random parameters' ex-post distribution. We finish by presenting the results from simulations that were performed in order to evaluate the impact of prices on CMP choice.

3.7.1 Random parameters ex-post distribution²⁷

First, we investigate the random parameters from the technology choice model *i.e.* η_c and μ_{cl} , and ρ_0 and ρ . Estimation standard errors as well as mean and standard deviation of the ex-post distribution of these random parameters are gathered in Table 3.2. Overall, estimation standard errors are small *i.e.* ex-post distribution of these random parameters benefit from precise estimates. η_c and μ_{cl} are respectively the cost parameters – economical and non-economical costs (*e.g.*, environmental concerns) – from the initial and transition probability functions. The reference CMP being the most intensive one (*i.e.* $\eta_1 \equiv 0$), η_2 and η_3 represent the relative cost of the intermediate and low-input technology in the initial probability function of CMP choice. On average, the intermediate CMP is less expensive than the high-yielding one as it has a negative sign. On the other hand, the low-input CMP appears as more expensive. We might think about the learning and opportunity costs associated to the low-input CMP.

Likewise, in the transition probability we set the most intensive CMP as reference and thus fixed $\mu_{c1} \equiv 0, \forall c \in C$. It implies that μ_{ck} corresponds to the cost to switch from CMP c to k relatively to a switch to the most intensive CMP. First, we can observe that it is systematically more expensive to switch to the low-input CMP (both μ_{13} and μ_{23} are positive). Yet, when adopted, the low-input CMP is meant to be stable as μ_{33} is negative *i.e.* staying in this CMP is the least costful option. As for the intermediate technology, we can observe that μ_{12} is negative, *i.e.* it is on average less expensive to switch from a high-yielding CMP to an intermediate one than to keep the high-yielding CMP. On the other hand μ_{22} is positive, meaning that changing to the high-input technology is less expensive than keeping the intermediate CMP. Such findings are a little bit surprising as one could expect that staying in the same CMP might be the dominant strategy. Yet, as mentioned previously, the switching costs are rather limited. Thus, random parameters μ_{ck} also capture the systematic cost differences between CMP. Then, one can argue that the high-yielding CMP has larger systematic costs than the intermediate one, hence the negative mean for parameter μ_{12} .

²⁷Plotted distribution from these random parameters can be found in Appendix 3.9.4.

TABLE 3.2: Estimation standard errors, mean and standard deviation from the ex-post distribution of the CMP choice probability functions random parameters

	Estimation se	Mean	sd
Scale parameters			
ρ_0	0.001	1.099	0.062
ρ	0.001	0.976	0.034
Cost parameters			
η_1	.	0	0
η_2	0.002	-0.039	0.134
η_3	0.002	0.038	0.106
μ_{11}	.	0	0
μ_{12}	0.003	-0.364	0.174
μ_{13}	0.002	0.238	0.178
μ_{21}	.	0	0
μ_{22}	0.003	0.382	0.288
μ_{23}	0.004	0.255	0.162
μ_{31}	.	0	0
μ_{32}	0.002	0.160	0.229
μ_{33}	0.003	-0.289	0.181

Note: se = standard error; sd = standard deviation.

Source: Authors' calculations on CDER data.

As for ρ_0 and ρ , they measure the size of the error term in the initial and transition probabilities of CMP choice. The higher these parameters are, the lower is the size of the error term and the more expected returns and costs play a big role in CMP choice probabilities. As $\rho_0 > \rho$ on average, latest factors tend to have a greater role in the initial probability than in the transition one. This means that technology change is more exposed to random events than the baseline technology choice. Meaning that, even if another technology would seem to be more profitable on a specific time period, a change in technology requires more than a temporary increase in profit. Yet, ρ_0 is more dispersed with a standard error twice as much as the one observed for ρ . It might mean that farmers' behavior is more homogeneous when considering a change in technology than when considering the initial technology choice. Otherwise, it can indicate that the model for transition probability is better adapted to data than the model for initial choice.²⁸ More generally, scale parameters are less dispersed than the cost parameters, *i.e.* farmers' heterogeneity is greater when considering technology costs.

Overall, what stands out from the study of the random parameters from the probability of CMP choice is the peculiarity of the low-input CMP. Indeed, there is no apparent interest in switching to the low-input technology when a farmer use the

²⁸In particular, we might consider initial probabilities that are time-dependent, as suggested before. It might provide a better adjustment to the data.

high-input or intermediate CMP. Plus, the initial cost for adopting the low-input CMP is higher than for the high-yielding and intermediate CMPs. One could thus wonder why any farmer would adopt this low-technology. Yet, when one has adopted low-input technology, they have no incentive to change technology. This can be explained that the adoption of low-input technology is mainly driven by non-economic consideration *i.e.* if the farmer does not have such environmental consideration, he has no interest in adopting such technology. Farmers who value greatly the ecological impacts of their production choices have no interest switching to more intensive technologies.

Last random parameters we need to explore are b^1 and a^c . Their respective mean and standard deviation are listed in Table 3.3. Parameters b^1 represent input use and output levels of the reference technology, *i.e.* the high-yielding one. Parameters a^c represent the discount when changing technology. In particular, a^2 (respectively a^3) represents the discount from the high-yielding CMP to the intermediate one (respectively from the intermediate to the low-input CMP). When considering the switch from the high-yielding CMP to the intermediate one, the greater discount is associated to pesticides with an average discount parameter around 9.5%. The mean discount for fertilizers is about 6% and is around 5% for yields. Except for pesticides, the mean discounts between the intermediate and the low-input CMP are greater and are all around 7%. A more global remark is that discount parameters a^c are rather concentrated, which argues in favor of the stability of the built classes.

TABLE 3.3: Estimation standard errors, mean and standard deviation from the ex-post distribution of random parameters from the yield and input use models

	Estimation se	Mean	sd
Input use & output levels			
b_{fert}^1	0.006	2.240	0.135
b_{pest}^1	0.002	1.957	0.299
b_{yield}^1	0.004	7.973	0.799
Discount parameters			
a_{fert}^2	0.003	0.937	0.012
a_{pest}^2	0.006	0.906	0.021
a_{yield}^2	0.005	0.952	0.009
a_{fert}^3	0.004	0.929	0.008
a_{pest}^3	0.005	0.929	0.010
a_{yield}^3	0.005	0.925	0.012

Note: se = standard error; sd = standard deviation.

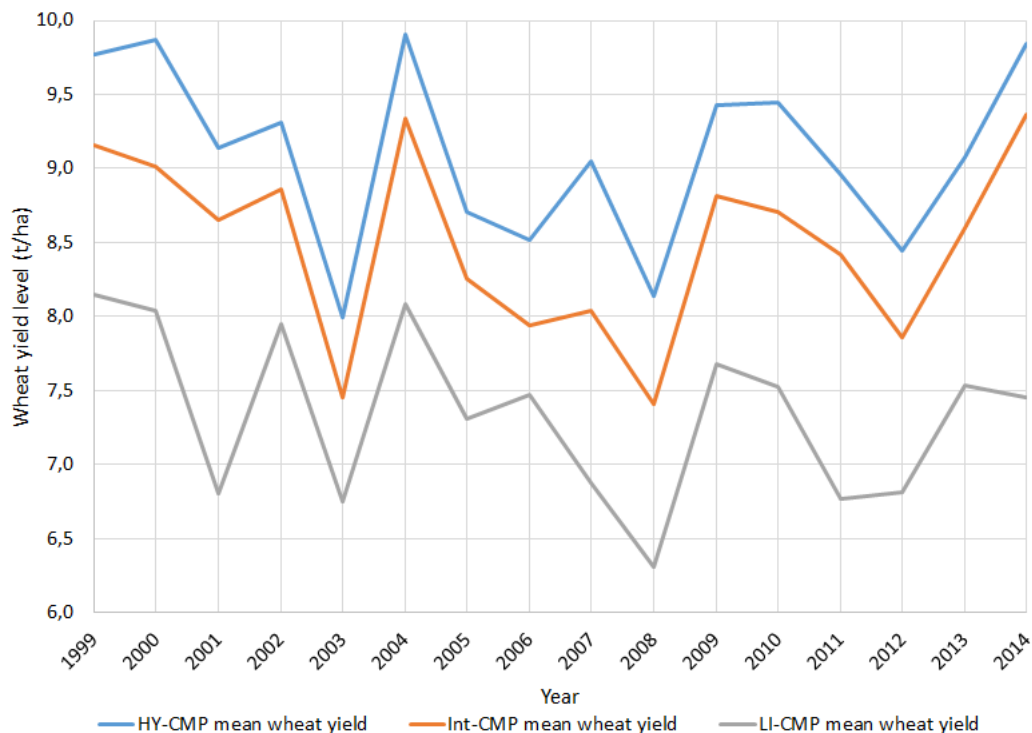
Source: Authors' calculations on CDER data.

More generally, from the review of the random parameters, we can say that our model (i) provides precise estimations of random parameters and (ii) most of them are coherent with what was expected.

3.7.2 Characteristics of the three CMP categories

Let now consider in more details the three categories that were distinguished thanks to the RPHMM: (i) high-input CMPs that are associated to larger levels of yield and input use (for both fertilizer and pesticide uses), (ii) intermediate CMPs with slightly lower yield and pesticide use levels but similar fertilizer use levels, (iii) low-input CMPs with lower yield and input use levels. The mean yield series from 1999 to 2014 – respectively the mean input uses series – observed in each of those three CMP categories are depicted in Figure 3.9 – respectively Figure 3.10.²⁹ From these two figures, we can see that pesticide use and yield levels characterize the three CMP categories. On average, intermediate CMPs are using 1.8% less nitrogen and 16.5% less pesticide than high-input CMPs for a 6.7% decrease in yields. Low-input CMPs use on average 3.4% (respectively 5.2%) less nitrogen, 15.4% (respectively 29.3%) less pesticides than intermediate CMPs (respectively high-input CMPs) for an average decrease of 13.4% (respectively 19.2%) in yields. These averages on the whole time period hide significant differences across years.

FIGURE 3.9: Annual mean yields for the three CMP categories obtained with RPHMM, from 1999 to 2014

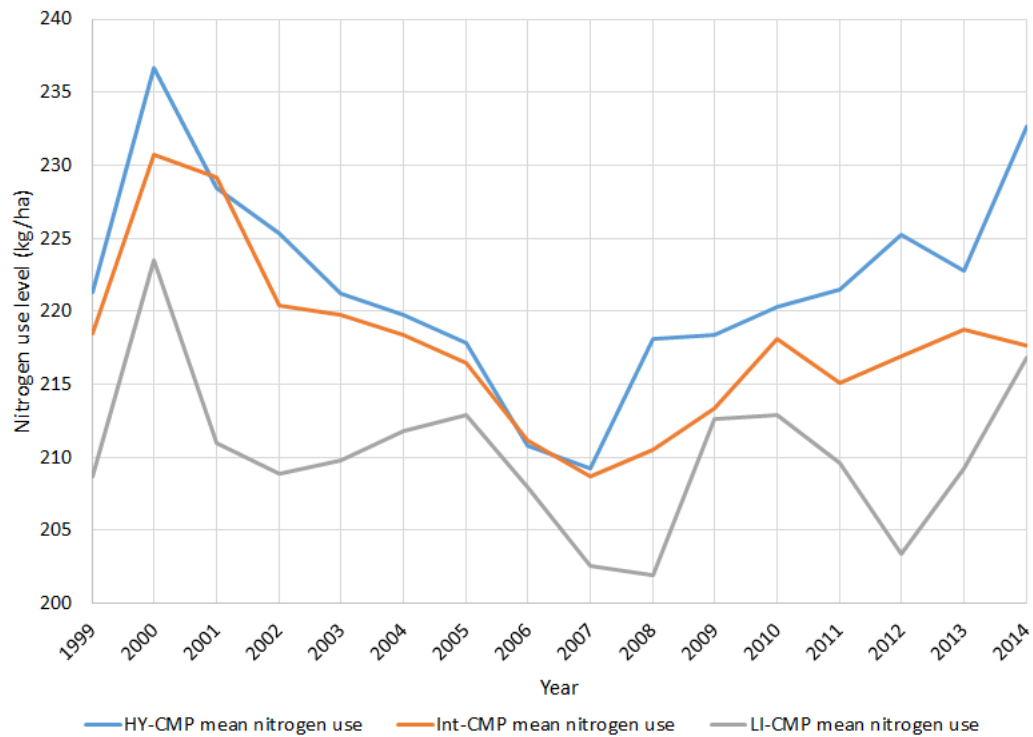


Source: Authors' calculations on CDER data.

²⁹Actually, it corresponds to $E[b_i^c] + d_t^0 + \delta_0' E[z_{it}|r]$.

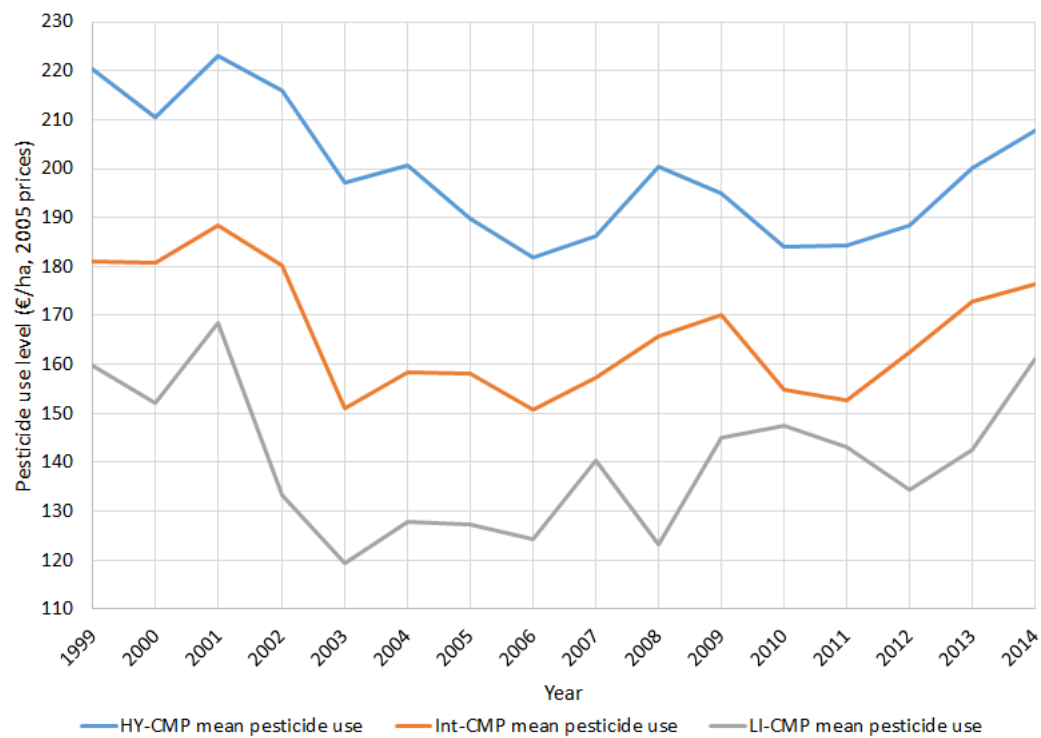
FIGURE 3.10: Annual mean input uses for the three CMP categories obtained with RPHMM, from 1999 to 2014

(A) Mean nitrogen use in each CMP type



Source: Authors' calculations on CDER data.

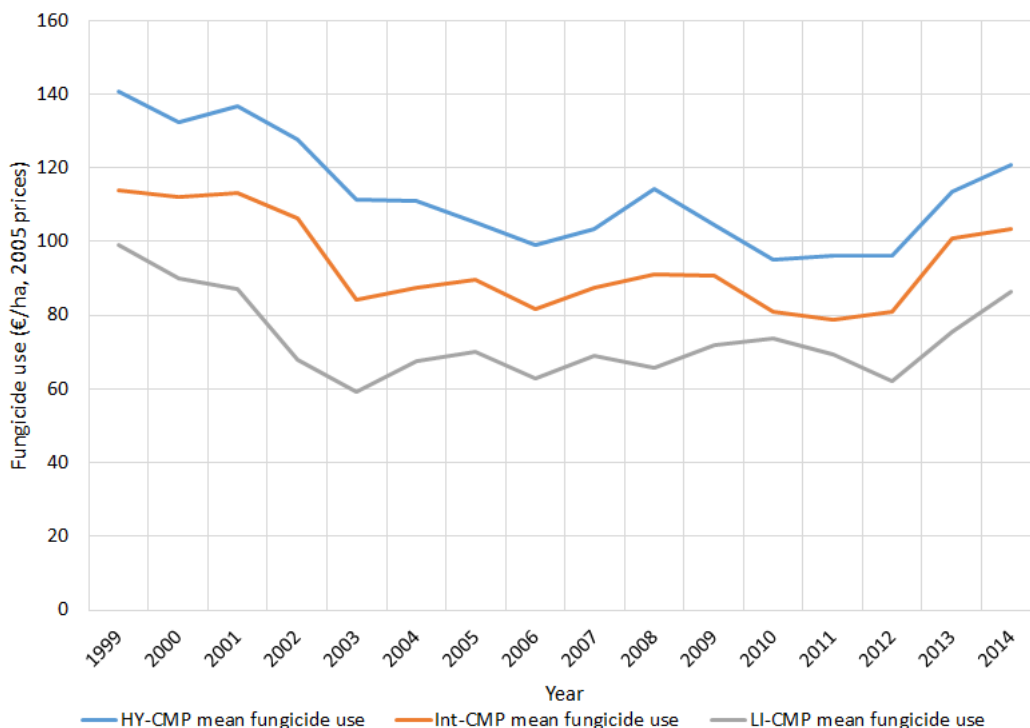
(B) Mean pesticide use in each CMP type



Source: Authors' calculations on CDER data.

In addition to overall pesticide uses, we investigate insecticide, fungicide and herbicide uses. From Figures 3.11 and 3.12, we can see that fungicide uses discriminate the three CMP categories. As for herbicides and insecticides, the discrimination is well-marked for high-input CMPs *versus* intermediate and low-input CMPs. Yet, uses among intermediate and low-input CMPs tend to overlap, implying a lower discrimination power of herbicides and insecticides than fungicides. A closer look at the relative differences between high- and low-input fungicide, herbicide and insecticide uses is also insightful. On average, fungicide and herbicide uses are 33% lower in low-input CMPs than in high-input ones. When targeting lower yields, farmers can use more resistant crop varieties and lower their sowing density, hence reducing their need in fungicides. As for herbicides, they can be – at least partially – substituted for by mechanical weeding. Differences in insecticide uses are smaller as on average, low-input CMPs use 21% less insecticides than high-input ones. Even if pest infestations can be partially avoided by a well-designed crop rotation scheme and more resistant crop varieties they cannot be totally avoided. And insecticides remain the most efficient way to get rid of them. This could explain why the difference is less marked for insecticide uses compared to herbicide and fungicide uses.

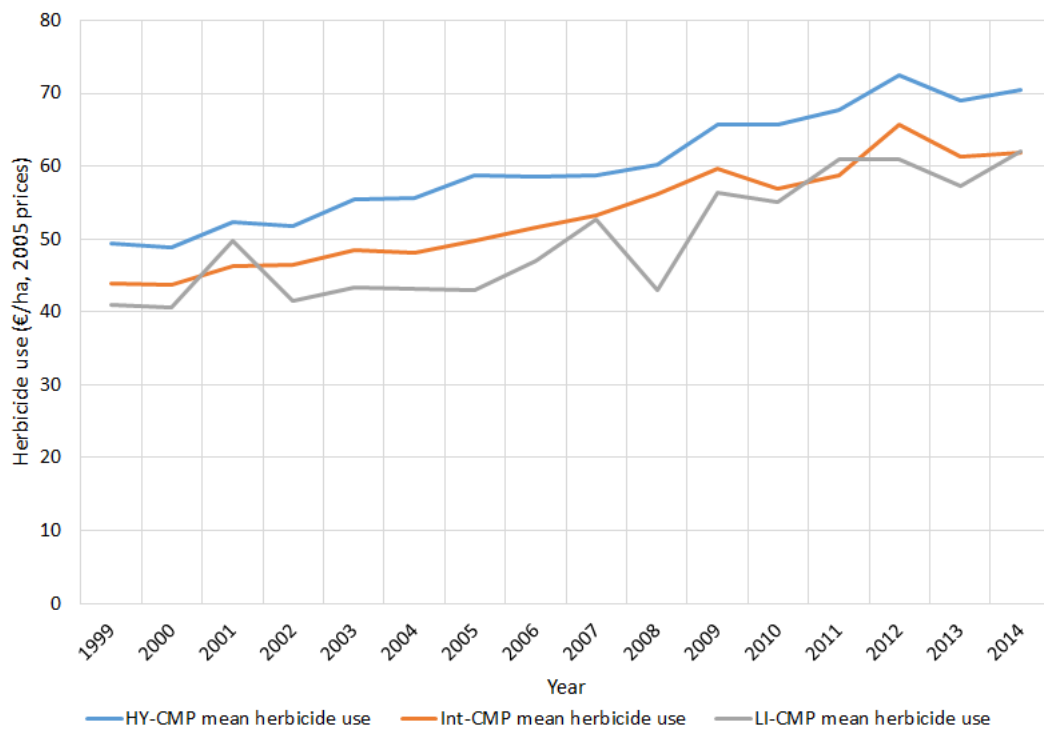
FIGURE 3.11: Annual mean fungicide uses for the three CMP categories obtained with RPHMM, from 1999 to 2014



Source: Authors' calculations on CDER data.

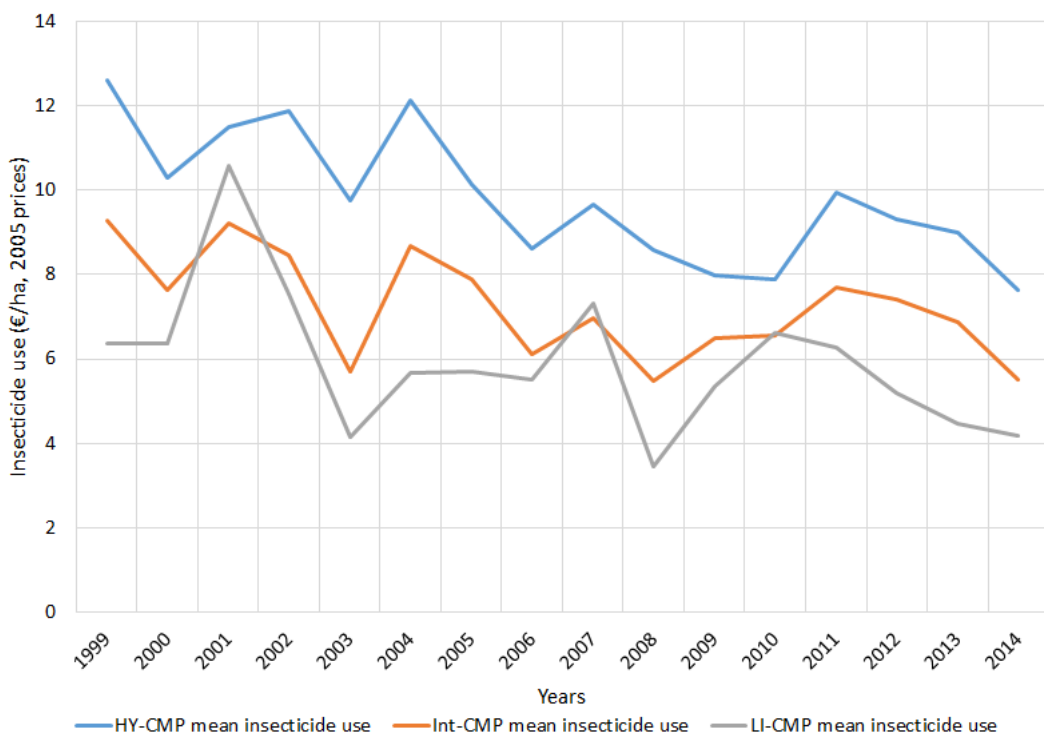
FIGURE 3.12: Annual mean herbicide uses for the three CMP categories obtained with RPHMM, from 1999 to 2014

(A) Mean herbicide use in each CMP category



Source: Authors' calculations on CDER data.

(B) Mean insecticide use in each CMP category

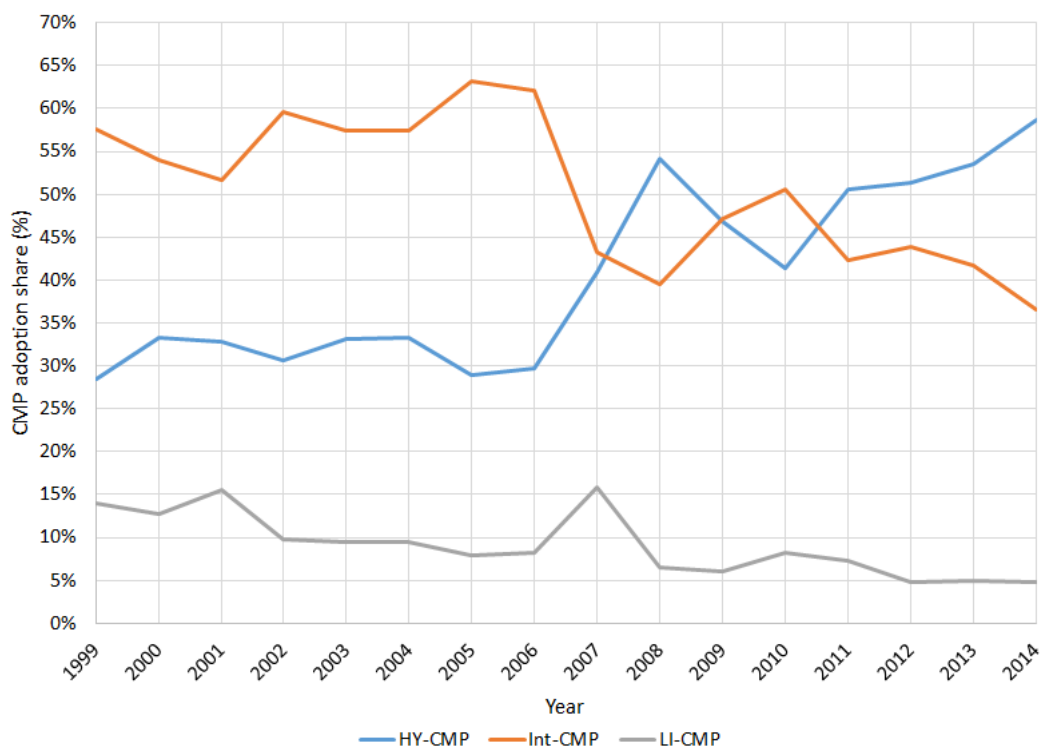


Source: Authors' calculations on CDER data.

Overall, these results tend to show that our RPHMM uncovers contrasted CMPs that are close to the “maximum yield”, “conventional” and “low-input/multi-resistant varieties” CMPs considered in Rolland et al. (2003) and Loyce et al. (2012). First, use of LI-CMPs reduces yield and fertilizer use levels by around 10% and pesticide use levels by around 30%. Second, the reduction in the use of pesticides is mostly due to fungicide uses.

Differences in input uses and in yields impact the expected profit associated to each CMP category. Figure 3.13 shows the evolution of the estimated share of farmers of our sample using each CMP category. Until 2007 crisis, these shares are steady with approximately 10% of farmers using low-input techniques, 60% using intermediate techniques and around 30% using high-input techniques. The estimated share of farmers using LI CMPs after 2007 is slightly inferior but is quite steady around 5%. Changes are observed among the shares of high-input and intermediate CMPs with a sharp increase in the use of high-input techniques in 2007/2008. These results tend to confirm the idea that high-input techniques are more profitable when wheat prices are high.

FIGURE 3.13: Estimated annual share of farmers who adopted a low-input CMP, from 1999 to 2014

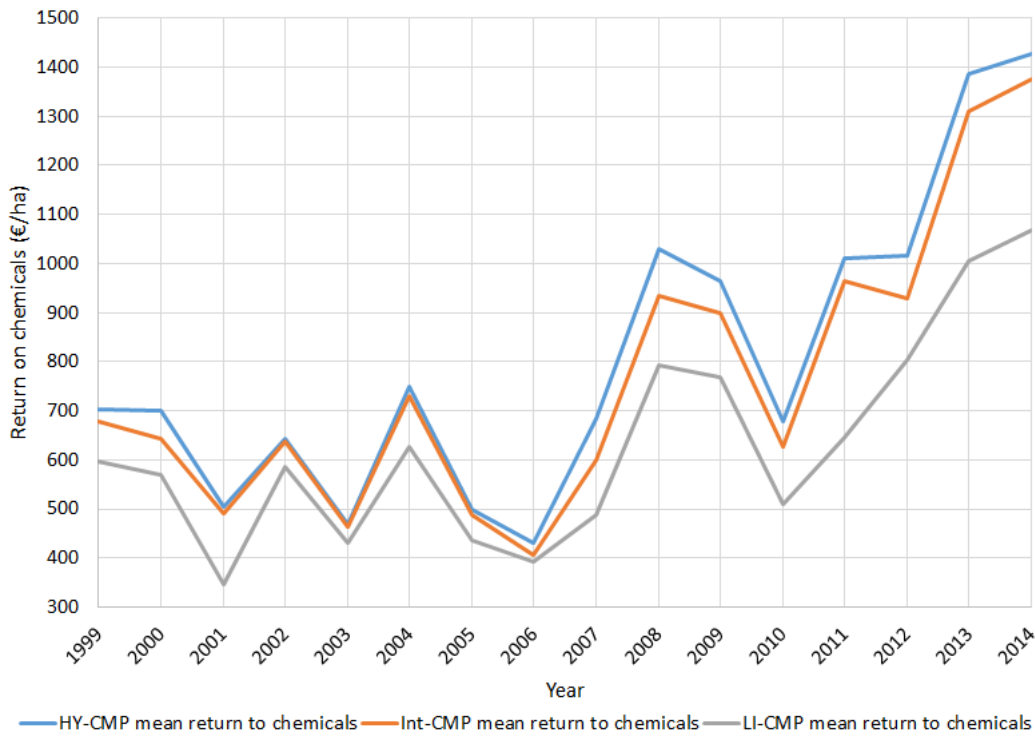


Source: Authors' calculations on CDER data.

Figure 3.14 displays that, as for the expected return associated to each CMP categories, 2007 is a pivotal year given that wheat prices suddenly increased in 2007 and remained relatively high, on average, since then. This increase in the price of wheat is associated with larger gaps in terms of expected return associated to each technology, explaining why high-input CMPs became more attractive. When wheat prices are lower, as in 2009, the gaps between expected returns of each CMP categories tend to be smaller and intermediate CMPs are more attractive to farmers. Figure 3.14 shows that expected returns are strongly linked to wheat prices with a 1-year delay (due to the fact we use wheat price of year $t - 1$ as anticipated price for year t). And higher prices tend to increase the differences between CMP technologies. The idea that low-input practices are more returnable in a context of low crop prices is also pointed out in the agronomic literature on CMPs (*e.g.*, Loyce et al., 2012; Rolland et al., 2003). Yet, in our case, the share of low-input practices is quite steady for the whole period. This argues in favor of the theory that choosing low-input practices is not only a choice that obey to economic reasons but is also linked to farmer personal values (*e.g.*, environmental and societal concern, see Frey and Stutzer, 2006; Howley et al., 2015; Mzoughi, 2011 for instance).

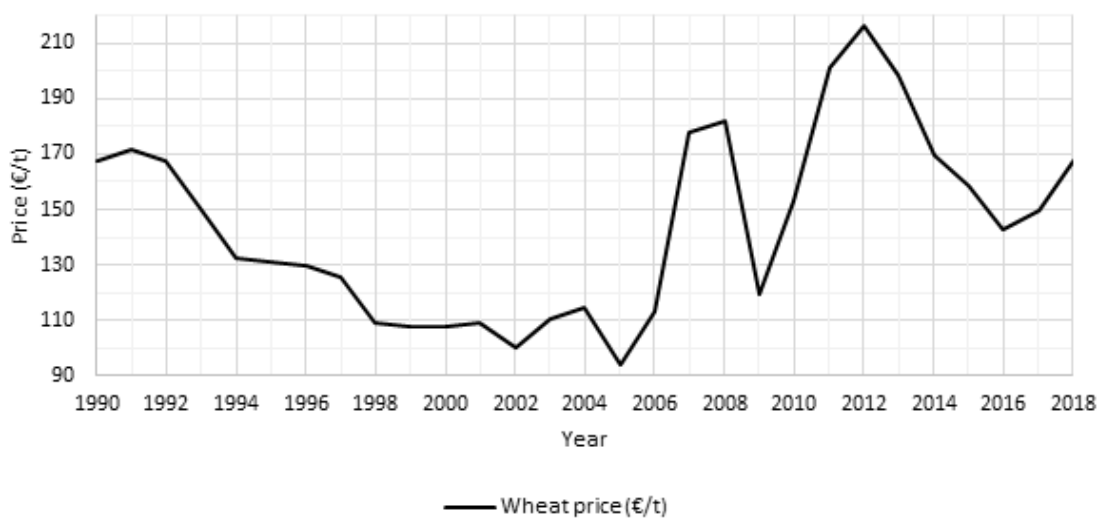
FIGURE 3.14: Estimated annual expected return for each CMP type (a) versus observed annual wheat prices (b)

(A) Annual expected return for the three CMP categories obtained with RPHMM, from 1999 to 2014



Source: Authors' calculations on CDER data.

(B) Annual wheat prices

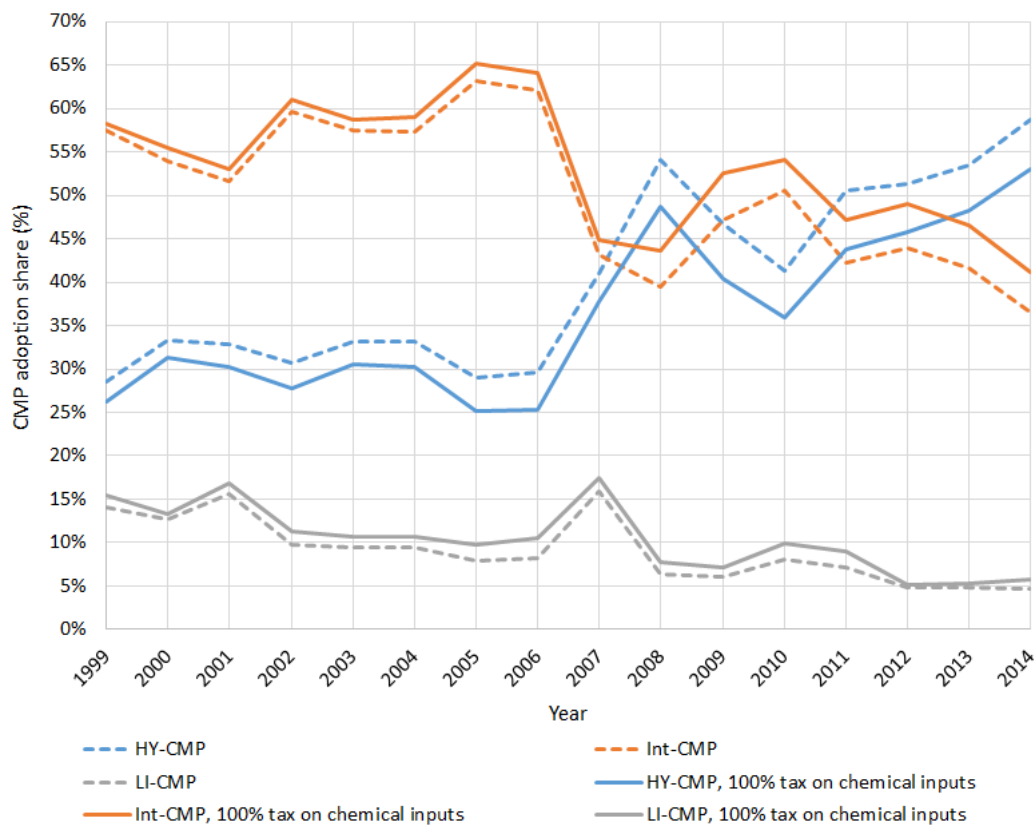


Source: RICA database.

3.7.3 Simulations results

Simulations were conducted as to investigate how farmers' CMP choice reacts to price changes. First, we simulate a 50% and 100% tax on chemical inputs. Figure 3.15 shows the respective size of each CMP categories at baseline and after simulating a 100% tax.³⁰

FIGURE 3.15: Annual change in the CMP adoption share after simulating a 100% tax on chemical inputs



Source: Authors' calculations on CDER data.

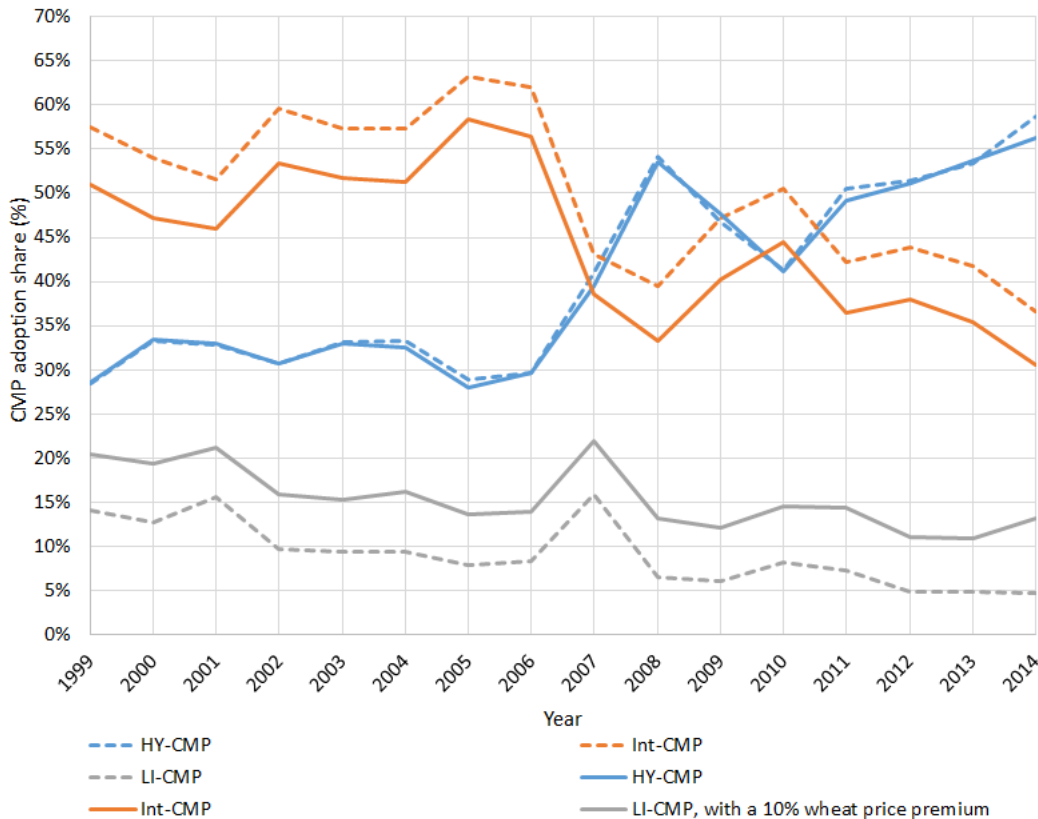
This 100% tax on chemical inputs entails a reduction in the high-input share in favor to the intermediate CMP. On average, the size of the high-input CMP decreases by 10.25% each year whereas the intermediate CMP size tends to increase by 6.30%. For the low-input CMP, the increase is even greater with a 15.70% annual increase on average. Yet, despite this rate, the low-input CMP remains marginal among wheat producers.

We also simulate differentiated wheat prices for low-input CMPs, *i.e.* low-input wheat producers benefit from a price premium. We consider three levels of price premium: a 5%, 10% and 20% price premium for low-input wheat prices. We present thereafter the results from the 10% price premium simulation.³¹

³⁰Results for the 50% tax simulation can be found in Appendix 3.9.5.

³¹Results for the 5% and 20% price premium can also be found in Appendix 3.9.5.

FIGURE 3.16: Annual change in the CMP adoption share after simulating a 10% price premium for low-input wheat producers



Source: Authors' calculations on CDER data.

In Figure 3.16 are the results from the 10% price premium. Compared to the tax simulation, the impact on the respective CMP adoption shares appears larger, especially from the low-input point of view. Indeed, with this 10% price premium, the share of farmers using the low-input CMP increases by more than 80% on average each year. On the contrary, the share of farmers using the high-yielding and intermediate CMP is barely impacted. When considering the 20% price premium, the size of the low-input CMP increases by more than 300% (annual average) whereas the high-yielding and intermediate CMP sizes decrease by respectively 6% and 51%. On that matter, we notice that the switches are mainly occurring from the intermediate to the low-input CMPs.

Overall, these simulations tend to show that the CMP adoption responsiveness to the tax instrument is rather limited. This is in line with previous studies on pesticides elasticity that tend to show that price elasticity of pesticide is low, *i.e.* price changes have little impact on pesticide use and on technology adoption to a larger extent. Regressing the relative prices of pesticides and fertilizers on the mean level on input use obtained from the model led to results demonstrating low responsiveness of these uses to price changes (or limited information content of our data for estimating

this responsiveness).³²

Our simulation results tend to show that chemical input savings associated to the low-input CMP are too small to compensate the revenue loss associated to lower yield levels. Even a 100% tax is insufficient to fill this gap. Price premiums tend to incentivize farmers to switch from intermediate to low-input CMPs. Thus, in a context of (i) high wheat prices, (ii) pesticide tax and (iii) in the absence of price premiums for low-input crops, it might be more appropriate from a policy perspective, to target farmers' switching directly to organic production as they would then benefit for a price premium on organic products.

3.8 Discussion and Conclusion

The random parameter hidden Markov model implemented in this article allowed us to identify three different cropping management practices among winter wheat producers of *La Marne*. Those cropping management practices are associated to different intensity levels that are designed to achieve different yield levels. Our modelling also allows to assess the effects of economic drivers on CMP choice. As we expected purely economic considerations, that is to say expected profit criteria, not to be the only drivers of CMP choices we accounted for possible effects of other, economic or non-economic, considerations in our model of farmers' CMP choice. These potential drivers of farmers' choice include unmeasured production costs, transition costs from a CMP to another or farmers' attitude toward risk or the environment. Not being able to disentangle the effects of this wide variety of CMP choice drivers limits our ability to analyze their effects in farmers' production choices and to provide insights on public policies aimed to foster the adoption of chemical input saving production practices by farmers. Yet, assessing the effects of purely economic drivers of CMP choices enables us to run simulations of public policies and to draw interesting conclusions regarding the efficiency of economic incentives, which is a unique feature of our micro-econometric modelling framework.

We proposed a random parameters model with endogenous regimes that follow a hidden Markov chain for uncovering the latent CMPs in a cost accounting panel dataset describing the production choices of a large sample of farmers. This model explicitly considers the latent CMP expected returns (to chemical inputs) as farmers' CMP choices. It is designed as a random parameters to account for farmer and farm unobserved heterogeneity. Farmers' CMP choice is defined as a Markov process to account for eventual CMP switching costs and farmers' possible reluctance to change their production practices. Our application on a panel dataset describing the wheat

³²Results from these regressions can be found in Appendix 3.9.6.

production choices of farmers located in the Marne area yields very interesting results.

First, this modelling framework allowed us to uncover three different CMP types used in the Marne area from 1999 to 2014. High-yielding CMPs are used by farmers seeking to achieve high yield levels. It relies on large nitrogen and pesticide uses. Intermediate CMPs allow to achieve slightly lower yield levels than those achieved based on the high-yielding CMPs. It also relies on slightly lower nitrogen and pesticide use levels. High-yielding and intermediate CMPs are standard, or conventional, crop production practices in France. By contrast, low-input CMPs are innovative production practices. Low-input practices were designed by agronomists for lowering chemical input uses, especially pesticide uses, by lowering the target yield levels (*e.g.*, Loyce et al., 2008, 2012; Meynard, 1991). Importantly, the characteristics of the low-input CMPs uncovered by our modelling framework are very close to those tested by agronomists in the Marne area during the mid 1990s (Larédo and Hocdé, 2014). Their yield and nitrogen use levels is about 10% lower than those of conventional CMPs, and their pesticide use levels are 30% lower. Also, our estimates reveal that most of the difference in pesticide uses is due to a reduction in fungicide uses in the low-input CMPs, which is consistent with the features of the low-input tested by agronomists in the Marne area (Loyce et al., 2008, 2012).

Second, the estimated model enable us to assess the expected returns of the considered CMPs, and their evolution during the considered period. The evolution of the differences in the CMP return is consistent with those of the adoption rates of the considered CMPs. In particular, the upward shift of wheat prices after 2006 led farmers to switch from intermediate CMPs to high-yielding CMPs and to switch from low-input CMPs to more intensive ones. Yet, the post 2007 wheat price levels significantly increased the differences in expected returns between the low-input CMPs and the other ones, with gaps ranging from 200 to 400 €/ha, but they did not fully deterred the use of low-input CMPs. This strongly suggests that non-economic motives impact farmers' production choices, at least those of some farmers (at least 5% in our case). Non-financial drivers of farmers' choices may include attitude toward the environment, health concerns and taste for agronomy and tend to play a great role in technology adoption (see, *e.g.*, Howley et al., 2015).

Lastly, the simulations we performed tend to show that input uses differences between low-input and more conventional CMPs are too small for taxes on chemical inputs to imply large relative profitability effects. This limited responsiveness of input uses to prices is in line with the literature showing low price elasticity of pesticides. Our finding that CMP choice is more responsive to a low-input price premium suggests that the decrease in expected yields implied by the use of low-input CMPs leads to reductions in revenues that cannot be compensated by the implied

savings in chemical input expenditures, especially when wheat prices are relatively high. This suggests in turn that taxes on chemical inputs may lead farmers to directly switch to organic production practices. The significant yield reduction induced by organic practices can be compensated by both significant “organic product” price premiums and large reductions in chemical input expenditures. In other words, low-input CMPs may entail too small chemical input savings and too large drops in expected yields for being a viable alternative to either conventional CMPs or organic production practices.

The specific features of the diffusion dynamics of agricultural production technologies are ignored in our modelling framework. Learning-by-doing and learning-by-others mechanisms are empirically documented by economists, especially in developing countries (*e.g.*, Chavas and Nauges, 2020; Foster and Rosenzweig, 2010; Marra and Pannell, 2003). These features are also often discussed by agronomists. The French pesticide use reduction program, the so-called EcoPhyto plan, launched the DEPHY farm network for fostering learning-by-others mechanisms and data collection on the characteristics and performances of pesticide saving practices. Even if some of these features are implicitly accounted for in our modelling framework, the model we consider largely overlook them. Indeed, the fact that the adoption of the considered technologies is unobserved in our case makes it particularly difficult to account for them. Yet, the adoption rate of low-input CMPs is likely to be rather limited in France, implying in turn that the effects of the congruent learning processes are likely to be limited as well.

We consider implementing some extensions to the actual modelling framework. First, in our Markov model the initial adoption probabilities of the considered CMPs are defined as functions of CMP expected returns and of time invariant (though farm specific) CMP specific costs. This raises specific issues in our application since the farmers joined (and quitted) our sample in various years. Farmers’ initial CMP choices in our data may depend on unobserved factors that vary across time, including unmeasured financial costs. For instance, dissemination of information on the implementation of low-input CMPs may lead to decreases in their implicit implementation costs. We are currently considering a version of our model that includes time trends in the probability functions of the initial CMP choice and in the transition probability functions.³³ Giving more structure to the latent yield and input use models considered in our modelling framework could also be fruitful, for further investigating these CMP specific production choices.

³³We are currently investigation such extension. Yet, the introduction of such time trend in the initial probability of adoption entails convergence issues.

We also consider applying our approach to other crops, which is possible with our dataset. Considering other crops is of particular interest as low-input CMPs have not been explicitly designed and promoted by agronomists for crops other than wheat. For instance, this would enable us to investigate if farmers using low-input CMPs for their wheat production extended the principles of the low-input CMPs to other crops. Taking a step further would lead us to multi-crop models, as in Koutchadé, Carpentier, and Féménia (2018, 2020). Our model is defined at the crop level while farmers also consider their other crops when choosing their production practices for a given crop, due to cropping systems effects and crop rotation effects in particular. Yet, considering a multi-crop framework with unobserved CMP choice appears to be particularly challenging.

3.9 Appendices

3.9.1 Detailed estimation procedure

SAEM algorithm explicit forms

Terms $E^{(n)}[\ln \ell_i^C(\boldsymbol{\theta}|\mathbf{r}_{(i)}, \boldsymbol{\alpha}_i)|\zeta_{(i)}]$ involve to compute expectations condition on the observed data $\zeta_{(i)}$. We use simulation methods to integrate those conditional expectations. The stochastic EM algorithm we use here is an extension of the SAEM algorithm proposed by Delyon, Lavielle, and Moulines (1999). It consists in dividing the E-step of the standard EM algorithm into three steps: a simulation step (S), an (intra) expectation step (E) and a stochastic approximation step (SA). We describe these steps in what follows.

At iteration n , given observed data $\zeta_{(i)}$ and the current value of parameter $\boldsymbol{\theta}$, $\boldsymbol{\theta}^{(n-1)}$:

1. **S-step:** Simulate $\left\{ \hat{\boldsymbol{\alpha}}_i^{(n,m)} : m = 1, \dots, R \right\}$ according to the conditional distribution $p\left(\boldsymbol{\alpha}_i \mid \zeta_{(i)}, \boldsymbol{\theta}^{(n-1)}\right)$, for $i = 1, \dots, N$
2. **E-step:** Given $\left\{ \hat{\boldsymbol{\alpha}}_i^{(n,m)} : m = 1, \dots, R \right\}$, evaluate the quantities $p_t(c \mid \mathbf{w}_{(i)}, \hat{\boldsymbol{\alpha}}_i^{(n,m)})$ and $s_t(c, d \mid \mathbf{w}_{(i)}, \hat{\boldsymbol{\alpha}}_i^{(n,m)})$, for $i = 1, \dots, N$, $t = t_{(i)}, \dots, T_{(i)}$, $c \in C$ and $d \in C$.
3. **SA-step:** update sufficient statistics according to

$$\left(\begin{array}{l} s_{it,c}^{(\zeta,n)} = s_{it,c}^{(\zeta,n-1)} + \lambda_{(n)} \left(R^{-1} \sum_{m=1}^R p_t(c \mid \mathbf{w}_{(i)}, \hat{\boldsymbol{\alpha}}_i^{(n,m)}) - s_{it,c}^{(\zeta,n-1)} \right) \\ s_{it,cd}^{(\xi,n)} = s_{it,cd}^{(\xi,n-1)} + \lambda_{(n)} \left(R^{-1} \sum_{m=1}^R s_t(c, d \mid \hat{\boldsymbol{\alpha}}_i^{(n,m)}, \mathbf{w}_{(i)}) - s_{it,cd}^{(\xi,n-1)} \right) \\ \mathbf{s}_i^{(\boldsymbol{\alpha},n)} = \mathbf{s}_i^{(\boldsymbol{\alpha},n-1)} + \lambda_{(n)} \left(R^{-1} \sum_{m=1}^R \hat{\boldsymbol{\alpha}}_i^{(n,m)} - \mathbf{s}_i^{(\boldsymbol{\alpha},n-1)} \right) \\ \mathbf{s}_i^{(\boldsymbol{\alpha}\boldsymbol{\alpha},n)} = \mathbf{s}_i^{(\boldsymbol{\alpha}\boldsymbol{\alpha},n-1)} + \lambda_{(n)} \left(R^{-1} \sum_{i=1}^N \sum_{m=1}^R \hat{\boldsymbol{\alpha}}_i^{(n,m)} \left(\hat{\boldsymbol{\alpha}}_i^{(n,m)} \right)' - \mathbf{s}_i^{(\boldsymbol{\alpha}\boldsymbol{\alpha},n-1)} \right) \\ \mathbf{s}_{it,c}^{(\boldsymbol{\varepsilon},n)} = \mathbf{s}_{it,c}^{(\boldsymbol{\varepsilon},n-1)} + \lambda_{(n)} \left(R^{-1} \sum_{r=1}^R p_t(c \mid \mathbf{w}_{(i)}, \hat{\boldsymbol{\alpha}}_i^{(n,m)}) (\mathbf{q}_{it} - \mathbf{b}_i^{c(n,m)}) - \mathbf{s}_{it,c}^{(\boldsymbol{\varepsilon},n-1)} \right) \\ \mathbf{s}_c^{(\boldsymbol{\varepsilon}\boldsymbol{\varepsilon},n)} = \mathbf{s}_c^{(\boldsymbol{\varepsilon}\boldsymbol{\varepsilon},n-1)} + \lambda_{(n)} \left(R^{-1} \sum_{i=1}^N \sum_{t=1}^T \sum_{m=1}^R \left[\begin{array}{l} p_t(c \mid \mathbf{w}_{(i)}, \hat{\boldsymbol{\alpha}}_i^{(n,m)}) (\mathbf{q}_{it} - \mathbf{b}_i^{c(n,m)}) \\ (\mathbf{q}_{it} - \mathbf{b}_i^{c(n,m)})' - \mathbf{s}_c^{(\boldsymbol{\varepsilon}\boldsymbol{\varepsilon},n-1)} \end{array} \right] \right) \end{array} \right).$$

From those three step we can deduce the approximation of $E^{(n)}[\ln \ell_i^C(\boldsymbol{\theta}|\mathbf{r}_{(i)}, \boldsymbol{\alpha}_i)|\zeta_{(i)}]$ given by:

$$\begin{aligned} & E^{(n)}[\ln L^C(\boldsymbol{\theta})|\zeta] \\ & \simeq \\ & \sum_{i=1}^N \left\{ \begin{array}{l} E^{(n)}[R(\boldsymbol{\alpha}_i, \mathbf{w}_{(i)})|\zeta_{(i)}] \\ + \sum_{t=t_{(i)}}^{T_{(i)}} \sum_{c \in C} \sum_{j=1}^{J^{(n)}} \hat{\omega}_i^{(n),j} s_t(c, d \mid \hat{\boldsymbol{\alpha}}_i^{(n),j}, \mathbf{w}_{(i)}) \ln \varphi(\mathbf{q}_{it} - \hat{\mathbf{b}}_i^{c(n),j} - \mathbf{d}_t - \boldsymbol{\Delta} \mathbf{z}_{it}; \boldsymbol{\Sigma}^c) \\ + \sum_{j=1}^{J^{(n)}} \hat{\omega}_i^{(n),j} \ln \varphi(\hat{\boldsymbol{\alpha}}_i^{(n),j} - \boldsymbol{\alpha}; \boldsymbol{\Omega}) \end{array} \right\}. \end{aligned}$$

Then, we can realize the last step, *i.e.* the **M-step** that consists in updating parameter θ according to:

$$\begin{pmatrix} \alpha_0^{(n)} &= N^{-1} \sum_{i=1}^N \mathbf{s}_i^{(\alpha, n)} \\ \mathbf{\Omega}^{(n)} &= N^{-1} \mathbf{s}_i^{(\alpha, n)} (\alpha_0^{(n)})' \\ \delta^{(n)} &= \left(\sum_{i=1}^N \sum_{t=t(i)}^{T(i)} \sum_{c \in C} s_{it,c}^{(\zeta, n)} \mathbf{z}'_{it} (\mathbf{\Sigma}^{c(n-1)})^{-1} \mathbf{z}_{it} \right)^{-1} \\ &\quad \times \sum_{i=1}^N \sum_{t=t(i)}^{T(i)} \sum_{c \in C} \mathbf{z}'_{it} (\mathbf{\Sigma}^{c(n-1)})^{-1} \mathbf{s}_{it,c}^{(\varepsilon, n)} \\ \mathbf{\Sigma}^{c(n)} &= \left(\sum_{i=1}^N \sum_{t=t(i)}^{T(i)} s_{it,c}^{(\zeta, n)} \right)^{-1} \begin{pmatrix} \mathbf{s}_c^{(\varepsilon, n)} - \sum_{i=1}^N \sum_{t=t(i)}^{T(i)} \mathbf{z}_{it} \delta^{(n)} (\mathbf{s}_{it,c}^{(\varepsilon, n)})' \\ - \sum_{i=1}^N \sum_{t=t(i)}^{T(i)} (\mathbf{z}_{it} \delta^{(n)} (\mathbf{s}_{it,c}^{(\varepsilon, n)})')' \\ + \sum_{i=1}^N \sum_{t=1}^T s_{it,c}^{(\zeta, n)} \mathbf{z}_{it} \delta^{(n)} (\mathbf{z}_{it} \delta^{(n)})' \end{pmatrix} \end{pmatrix}.$$

Decreasing positive sequence

$\{\lambda_{(n)}\}$ sequence from (SA) step must be a decreasing positive sequence such that (i) $\lambda_{(1)} = 1$, (ii) $\sum_{n=1}^{+\infty} \lambda_{(n)} = +\infty$ and (iii) $\sum_{n=1}^{+\infty} \lambda_{(n)}^2 < +\infty$. This sequence defines the step of the stochastic approximation, impacts the speed of convergence as well as the algorithm's convergence to the ML. Kuhn and Lavielle (2005) proposes to set $\lambda_{(n)} = 1$ for the first n_1 iterations and then gradually reduce $\lambda_{(n)}$. We set here:

$$\lambda_{(n)} = \begin{cases} 1 & \text{for } 1 \leq n \leq n_1 \\ (n - n_1 + 1)^{3/4} & \text{for } n > n_1 \end{cases},$$

and n_1 is chosen very large to guarantee that the algorithm reaches the neighborhood of the ML before $\lambda_{(n)}$ starts to decrease.

Simulation step procedure

To perform (S) step at iteration n , we use a few Markov chain Monte-Carlo (MCMC) iterations with $p(\alpha|\zeta_{(i)}, \theta^{(n-1)})$ as stationary distribution, and we retain R MCMC draws for each i , $i = 1, \dots, N$.³⁴ We use Metropolis-Hastings (MH) algorithm with a random walk proposal distribution to simulate the chain with length $R + R_{\text{burn}}$ draws, *i.e.* we draw $\hat{\alpha}_{i,m}$ such that $\hat{\alpha}_{i,m} \sim N(\hat{\alpha}_{i,m-1}, \mathbf{\Psi})$ for $1 \leq i \leq N$ and $1 \leq m \leq R + R_{\text{burn}}$. We defined the acceptance rate as:

$$\tau(\hat{\alpha}_{i,m-1}, \hat{\alpha}_{i,m}) = \min \left(1, \frac{p(\mathbf{q}_{(i)}, \hat{\alpha}_{i,m} | \theta^{(n-1)})}{p(\mathbf{q}_{(i)}, \hat{\alpha}_{i,m-1} | \theta^{(n-1)})} \right).$$

Diagonal matrix $\mathbf{\Psi}$ is adaptively adjusted such as $\tau(\hat{\alpha}_{i,m-1}, \hat{\alpha}_{i,m}) \in [0.24, 0.40]$. After $R + R_{\text{burn}}$ iterations, the first R_{burn} draws are discarded as burn-ins and we only consider the last R draws.

³⁴In our case, we set $R = 1$ as we have many individuals.

Stopping rule of the algorithm and diagnostic plots

As in Koutchadé, Carpentier, and Féménia (2018, 2020) we use a standard stopping rule based on the relative changes in the values of the estimated parameters between two iterations (*e.g.*, Booth and Hobert, 1999; Booth, Hobert, and Jank, 2001). If the condition

$$\max_j \left(\frac{|\theta_{j,n} - \theta_{j,n-1}|}{|\theta_{j,n}| + \sigma_1} \right) < \sigma_2, \quad (3.15)$$

holds for three consecutive iterations for chosen positive values of convergence parameters θ_1 and θ_2 , the algorithm stops. In our case, we set up $\theta_1 = 0.01$ and $\theta_2 = 0.0001$. To ensure that parameters θ_n achieved, at least approximately, the maximum of the considered likelihood function when the condition (3.15) is met, we implement three safeguards. First, we implement this stopping rule only once we have reached an iteration index greater than n_1 (*cf.* the part on $\{\lambda_{(n)}\}$ sequence). Second, we check that the scores are null and that the Hessian matrix is negative definite at the estimated value of θ (Gu and Zhu, 2001). Third, we check graphically the convergence of parameters by plotting their values along iterations.

Estimation of the variance of the estimates

To estimate the variance of the estimated parameters θ , we use the procedure described by Ruud (1991). We use the MH algorithm to draw the sequence $\{\hat{\alpha}_{i,r} : r = 1, \dots, R\}$ from $p(\alpha | \zeta_{(i)}, \hat{\theta})$, for $i = 1, \dots, N$, where $\hat{\theta}$ are the estimates we obtained from the SAEM algorithm. Then, we can approximate the information matrix $I(\hat{\theta})$ by:

$$\tilde{I}(\hat{\theta}) = N^{-1} \sum_{i=1}^N \left(R^{-1} \sum_{r=1}^R \partial_{\theta} \log p(\mathbf{q}_{(i)}, \hat{\alpha}_{i,r}; \hat{\theta}) \right) \left(R^{-1} \sum_{r=1}^R \partial \log p(\mathbf{q}_{(i)}, \hat{\alpha}_{i,r}; \hat{\theta}) \right)',$$

and the variance of estimates by:

$$V(\hat{\theta}) = \tilde{I}(\hat{\theta})^{-1}.$$

Estimation of the likelihood and model selection

To estimate the likelihood and select the model we rely on *Monolix Methodology* (2014). Given the estimate $\hat{\theta}$ of θ , the log-likelihood of the model is given by:

$$\ell(\hat{\theta}) = \sum_{i=1}^N \log p(\mathbf{q}_{(i)}; \hat{\theta}),$$

with $p(\mathbf{q}_{(i)}; \hat{\theta}) = \int p(\mathbf{q}_{(i)}, \alpha_i; \hat{\theta}) d\alpha_i = \int p(\mathbf{q}_{(i)} | \alpha_i; \hat{\theta}) \varphi(\alpha_i - \hat{\alpha}; \hat{\Omega}) d\alpha_i$.

$p(\mathbf{q}_{(i)}; \hat{\theta})$ has no closed form, so we use the importance-sampling approach to estimate it. From prior distribution $\varphi(\alpha_i - \hat{\alpha}; \hat{\Omega})$ as importance density, we draw independence

sequence $\{\boldsymbol{\alpha}_{i,r} : r = 1, \dots, R\}$ and then approximate $p(\mathbf{q}_{(i)}; \hat{\boldsymbol{\theta}})$ by

$$p(\mathbf{q}_{(i)}; \hat{\boldsymbol{\theta}}) \simeq R^{-1} \sum_{r=1}^R p(\mathbf{q}_{(i)} | \hat{\boldsymbol{\alpha}}_{i,r}; \hat{\boldsymbol{\theta}}),$$

where $p(\mathbf{q}_{(i)} | \hat{\boldsymbol{\alpha}}_{i,r}; \hat{\boldsymbol{\theta}})$ is obtained using the Forward-Backward algorithm.³⁵ This estimator is unbiased and consistent as its variance decreases as $1/R$.

We also define the $-2LL$, AIC and BIC criteria as:

$$\begin{cases} -2LL & = -2l(\hat{\boldsymbol{\theta}}) \\ \text{AIC} & = 2l(\hat{\boldsymbol{\theta}}) + 2P \\ \text{BIC} & = 2l(\hat{\boldsymbol{\theta}}) + \log(N)P \end{cases},$$

where P is the total number of parameters to be estimated and N the number of observations.

Forward-Backward algorithm

Let start by defining the forward variable $\omega_{it}(c, \boldsymbol{\alpha}_i)$:

$$\omega_{it}(c, \boldsymbol{\alpha}_i) = p(\mathbf{q}_{it(i)}, \dots, \mathbf{q}_{it}, r_{it} = c | \boldsymbol{\alpha}_i),$$

for $i = 1, \dots, N$, $t = t_{(i)}, \dots, T_{(i)}$ and $c \in \mathcal{C}$. $\omega_{it}(c, \boldsymbol{\alpha}_i)$ denotes the probability for individual i to adopt CMP c at time t after seeing the partial sequence $(\mathbf{q}_{it(i)}, \dots, q_{it})$ given the random parameter $\boldsymbol{\alpha}_i$.

We can show that (see Maruotti, 2011):

$$\ell_i(\boldsymbol{\theta} | \boldsymbol{\alpha}_i) = \sum_{c \in \mathcal{C}} \omega_{iT_{(i)}}(c, \boldsymbol{\alpha}_i).$$

Terms $\omega_{it}(c, \boldsymbol{\alpha}_i)$ can be computed iteratively. Thus, $\omega_{it}(c, \boldsymbol{\alpha}_i)$ is given by:

$$\begin{cases} \omega_{it_{(i)}}(c, \boldsymbol{\alpha}_i) & = p_0(c | \boldsymbol{\alpha}_i, \mathbf{w}_{t_{(i)}}) \varphi(\mathbf{q}_{it(i)} - \mathbf{b}_i^c - \mathbf{d}_{t(i)} - \boldsymbol{\Delta} \mathbf{z}_{it(i)}; \boldsymbol{\Sigma}^c) \\ \omega_{it+1}(c, \boldsymbol{\alpha}_i) & = \sum_{l \in \mathcal{C}} \omega_{it}(l, \boldsymbol{\alpha}_i) p_{t+1}(c | l, \boldsymbol{\alpha}_i, \mathbf{w}_{t+1}) \varphi(\mathbf{q}_{it+1} - \mathbf{b}_i^c - \mathbf{d}_{t+1} - \boldsymbol{\Delta} \mathbf{z}_{it+1}; \boldsymbol{\Sigma}^c) \end{cases}.$$

Now, let start by defining the backward variable $\beta_{it}(c, \boldsymbol{\alpha}_i)$:

$$\beta_{it}(c, \boldsymbol{\alpha}_i) = p(\mathbf{q}_{it+1}, \dots, \mathbf{q}_{iT_{(i)}} | r_{it} = c, \boldsymbol{\alpha}_i),$$

for $i = 1, \dots, N$, $t = t_{(i)}, \dots, T_{(i)}$ and $c \in \mathcal{C}$. $\beta_{it}(c, \boldsymbol{\alpha}_i)$ denotes the probability of the partial sequence $(\mathbf{q}_{it+1}, \dots, \mathbf{q}_{iT_{(i)}})$ given that farmer i chooses CMP c at time t and

³⁵We present the Forward-Backward algorithm at the end of this Appendix.

given the random parameter α_i . We can compute this term iteratively by:

$$\begin{cases} \beta_{iT(i)}(c, \alpha_i) & = 1 \\ \beta_{it}(c, \alpha_i) & = \sum_{l \in C} p_{t+1}(l|c, \alpha_i, \mathbf{w}_{t+1}) \varphi(\mathbf{q}_{it+1} - \mathbf{b}_i^l - \mathbf{d}_{t+1} - \Delta \mathbf{z}_{it+1}; \Sigma^l) \beta_{it+1}(l, \alpha_i) \end{cases} .$$

We can show that, using the forward variable $\omega_{it}(c, \alpha_i)$, we have (see Maruotti, 2011):

$$\begin{cases} p_t(c|\mathbf{w}^{(i)}, \alpha_i) & = \frac{\omega_{it}(c, \alpha_i) \beta_{it}(c, \alpha_i)}{\sum_{c \in C} \omega_{it}(c, \alpha_i) \beta_{it}(c, \alpha_i)} \\ s_t(c, l|\alpha_i, \mathbf{w}^{(i)}) & = \frac{\omega_{it-1}(c, \alpha_i) p_t(l|c, \alpha_i, \mathbf{w}_t) \varphi(\mathbf{q}_{it} - \mathbf{b}_i^l - \mathbf{d}_t - \Delta \mathbf{z}_{it}; \Sigma^l) \beta_{it}(l, \alpha_i)}{\sum_{c \in C} \sum_{l \in C} \omega_{it-1}(c, \alpha_i) p_t(l|c, \alpha_i) \varphi(\mathbf{q}_{it} - \mathbf{b}_i^l - \mathbf{d}_t - \Delta \mathbf{z}_{it}; \Sigma^l) \beta_{it}(l, \alpha_i)} \end{cases} .$$

We also show that:

$$\begin{aligned} \ell_i(\theta|\alpha_i) &= \sum_{c \in C} \omega_{it}(c, \alpha_i) \beta_{it}(c, \alpha_i) \\ &= \sum_{c \in C} \sum_{l \in C} \omega_{it-1}(c, \alpha_i) p_t(l|c, \alpha_i) \varphi(\mathbf{q}_{it} - \mathbf{b}_i^l - \mathbf{d}_t - \Delta \mathbf{z}_{it}; \Sigma^l) \beta_{it}(l, \alpha_i) \end{aligned}$$

Estimation of the individual parameters and sequences of CMP

Given the estimate $\hat{\theta}$ of θ computed with the SAEM algorithm, we estimate individual parameters $\hat{\alpha}_i$ and the CMP sequence $\mathbf{r}_{(i)}$ using the two-step procedure as in Delattre and Lavielle, 2012. We make use of the condition distribution $p(\mathbf{r}_{(i)}, \alpha_i|\zeta_{(i)}, \hat{\theta})$ to first estimate parameters $\hat{\alpha}_i$ with the Maximum A Posteriori (MAP) approach:

$$\begin{aligned} \hat{\alpha}_i &= \arg \max_{\alpha_i} p(\alpha_i|\zeta_{(i)}, \hat{\theta}) \\ &= \arg \max_{\alpha_i} p(\zeta_{(i)}|\alpha_i, \hat{\theta}) p(\alpha_i|\hat{\theta}), \end{aligned}$$

since $p(\alpha_i|\zeta_{(i)}, \hat{\theta}) \propto p(\zeta_{(i)}|\alpha_i, \hat{\theta}) p(\alpha_i|\hat{\theta})$. We use R package *optim* to numerically optimize $\hat{\alpha}_i$.

We use the same MAP approach to estimate the unknown CMP sequence as

$$\mathbf{r}_{(i)} = \arg \max_{\mathbf{r}_{(i)}} p(\mathbf{r}_{(i)}|\zeta_{(i)}, \hat{\alpha}_i, \hat{\theta}).$$

Yet, to compute this equation, we need to use the Viterbi algorithm (Rabiner, 1989).³⁶

Viterbi algorithm

Let start by defining the following Viterbi path probability:

$$v_{it}(c, \alpha_i) = \max_{r_{it(i)}, \dots, r_{it-1}} p(r_{it(i)}, \dots, r_{it-1}, \mathbf{q}_{it(i)}, \dots, \mathbf{q}_{it}, r_{it} = c|\alpha_i).$$

³⁶As for the Forward-Backward algorithm, it is presented below.

This term can be computed iteratively by:

$$\begin{cases} v_{it(i)}(c, \boldsymbol{\alpha}_i) &= p_0(c|\boldsymbol{\alpha}_i, \mathbf{w}_{t(i)})\varphi(\mathbf{q}_{it(i)} - \mathbf{b}_i^c - \mathbf{d}_{t(i)} - \Delta\mathbf{z}_{it(i)}; \boldsymbol{\Sigma}^c) \\ v_{it+1}(c, \boldsymbol{\alpha}_i) &= \max_{l \in \mathcal{C}} v_{it}(l, \boldsymbol{\alpha}_i) p_{t+1}(c|l, \boldsymbol{\alpha}_i, \mathbf{w}_{t+1})\varphi(\mathbf{q}_{it+1} - \mathbf{b}_i^c - \mathbf{d}_{t+1} - \Delta\mathbf{z}_{it+1}; \boldsymbol{\Sigma}^c) \end{cases} \cdot$$

As taken from Rabiner, 1989, the best path of CMP, r_{it}^* , $t = t(i), \dots, T(i)$, can be found recursively by:

$$\begin{cases} r_{iT(i)}^* &= \arg \max_{c \in \mathcal{C}} v_{iT(i)}(c, \boldsymbol{\alpha}_i) \\ r_{it}^* &= \arg \max_{c \in \mathcal{C}} v_{it+1}(c, \boldsymbol{\alpha}_i) p_{t+1}(r_{it+1}^* | c, \boldsymbol{\alpha}_i, \mathbf{w}_{t+1}) \end{cases} \cdot$$

3.9.2 Insights about k-means and AHC algorithms

As said previously, clustering techniques allow to uncover groups among a data set. Two major elements to define in a clustering analysis are (i) the distance metric and (ii) the type of algorithm one wants to use. Indeed, there exists two main types of algorithms: partitioning and hierarchical algorithms.

Partitioning algorithms rely on the iterative reallocation of each observation to its closest cluster and k-means is amongst the most popular partitioning algorithms. There are two main ways to describe the k-means algorithm, depending on whether we use the one defined by MacQueen et al. (1967) or the one defined by Hartigan and Wong (1979). Both require to define the number of clusters we want to partition the data into and the distance metric to use *a priori*. Assuming the number of clusters to be k , step 0 of both algorithms is to define k initial centroids (*i.e.* the cluster “typical” observation). Then, MacQueen and Hartigan & Wong algorithms tend to differ slightly. Once k initial centroids have been assigned, at each iteration MacQueen algorithm assigns the observations to the closest cluster. Once this assignation done, clusters’ centroids are recalculated. The algorithm keeps going as long as (i) observations are not stabilized within a cluster or (ii) when the changes between centroids are still above a preliminary defined tolerance criterion.

As for Hartigan & Wong algorithm, a few differences can be highlighted even if the overall principle is similar. At step 0, once the k centroids are defined and before starting the iteration process, each observation is assigned to the closest centroid and clusters’ centroid are recalculated. Then, the reallocation during the iteration process relies on the notion of the within-cluster sum of squares. For each cluster whose centroid has changed between iteration $n - 1$ and n , we compute the within-cluster sum of squares. Then, for each observation among this cluster, we calculate the within-cluster sum of squares of the remaining clusters when including this observation. If we find that the within-cluster sum of squares of another cluster is inferior to the current one when including this particular observation, we reallocate it to this new cluster. Let consider the cluster 1 and the observation x_{i1} within this cluster. x_{i1} will be reallocated to another cluster j if:

$$SSE_j = \frac{N_j \sum_{i=1}^{N_j} \|x_{ij} - c_j\|^2}{N_j - 1} < \frac{N_1 \sum_{i=1}^{N_1} \|x_{i1} - c_1\|^2}{N_1 - 1} = SSE_1.$$

Algorithm keeps going as long as there are observations that keep changing cluster.

Other than partitioning algorithms are hierarchical algorithms. Hierarchical algorithms can either be ascending or descending. We only describe the ascending approach there as the descending approach can easily be deduced from that. As for

the k-means algorithm, a distance metric needs to be defined *a priori*. At step 0, each observation is considered as a cluster. Then, step 1 consists in (i) calculating the distance between each observations and (ii) gathering the two closest observations. The algorithm keeps going this way – except that instead of observations, we calculate the distance between clusters – until the whole dataset is gathered in a unique cluster. The number of clusters to keep can then be defined either *a priori* by the researcher or a statistical criteria can be used to define the optimal number of clusters to keep.³⁷

Unlike for k-means, hierarchical clustering results are not sensitive to the initialization step, as the initial step consists either in the whole dataset or each observation being a cluster. Yet, time complexity of hierarchical clustering increases quickly with the size of the data set as time complexity equals $O(n^2)$. A solution is to reduce data dimensionality (or “denoise” data) by performing a preliminary principal component analysis (PCA) and perform the ascendant hierarchical clustering on the obtained components (Husson, Josse, and Pagès, 2010).

3.9.3 Additional results for the "exploratory analyses"

First, we present the results from the latent class model using k-means results as initial values (instead of AHC).

³⁷This can also be done with k-means algorithms by comparing results obtained with different values of k .

FIGURE 3.17: Latent class model with k-means initials – Annual mean yield in high-yielding and low-input CMPs



Source: Authors' calculations on CDER data.

FIGURE 3.18: Latent class model with k-means initials – Annual mean input uses in high-yielding and low-input CMPs

(A) Annual mean nitrogen uses in high-yielding and low-input CMPs



Source: Authors' calculations on CDER data.

(B) Annual mean pesticide uses in high-yielding and low-input CMPs



Source: Authors' calculations on CDER data.

Then, we present the results we obtained from the clustering analyses meaning (i) AHC and (ii) k-means.

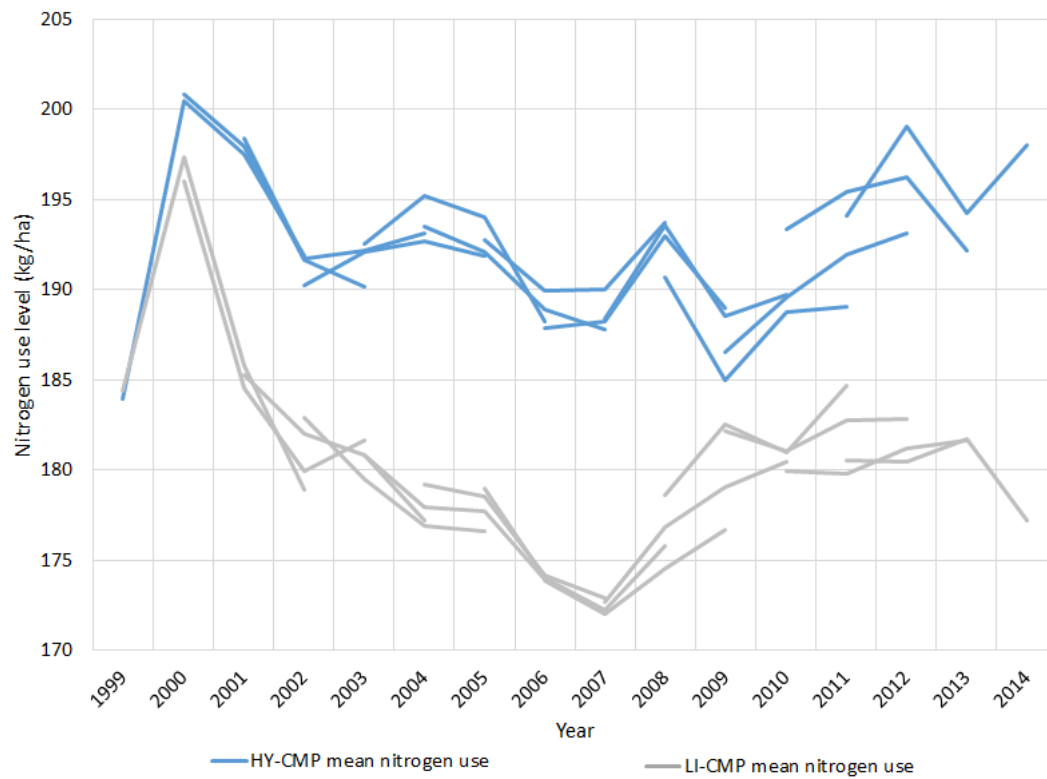
FIGURE 3.19: AHC clustering – Annual mean yield in high-yielding and low-input CMPs



Source: Authors' calculations on CDER data.

FIGURE 3.20: AHC clustering – Annual mean input uses in high-yielding and low-input CMPs

(A) Mean nitrogen uses in high-yielding and low-input CMPs



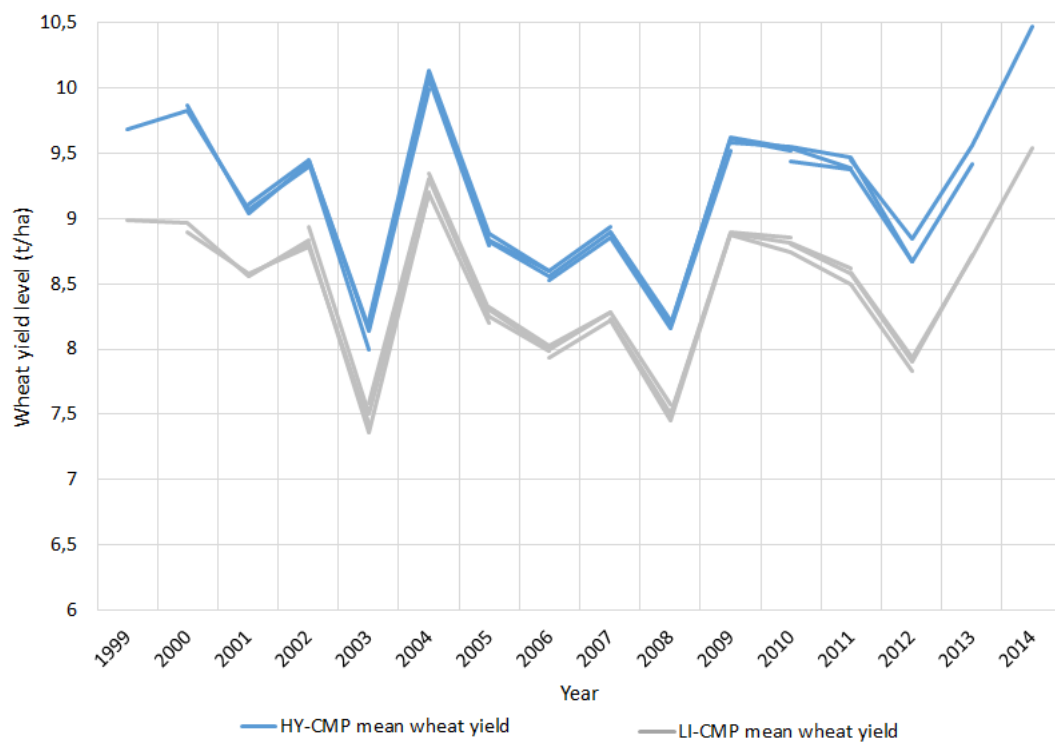
Source: Authors' calculations on CDER data.

(B) Mean pesticide uses in high-yielding and low-input CMPs



Source: Authors' calculations on CDER data.

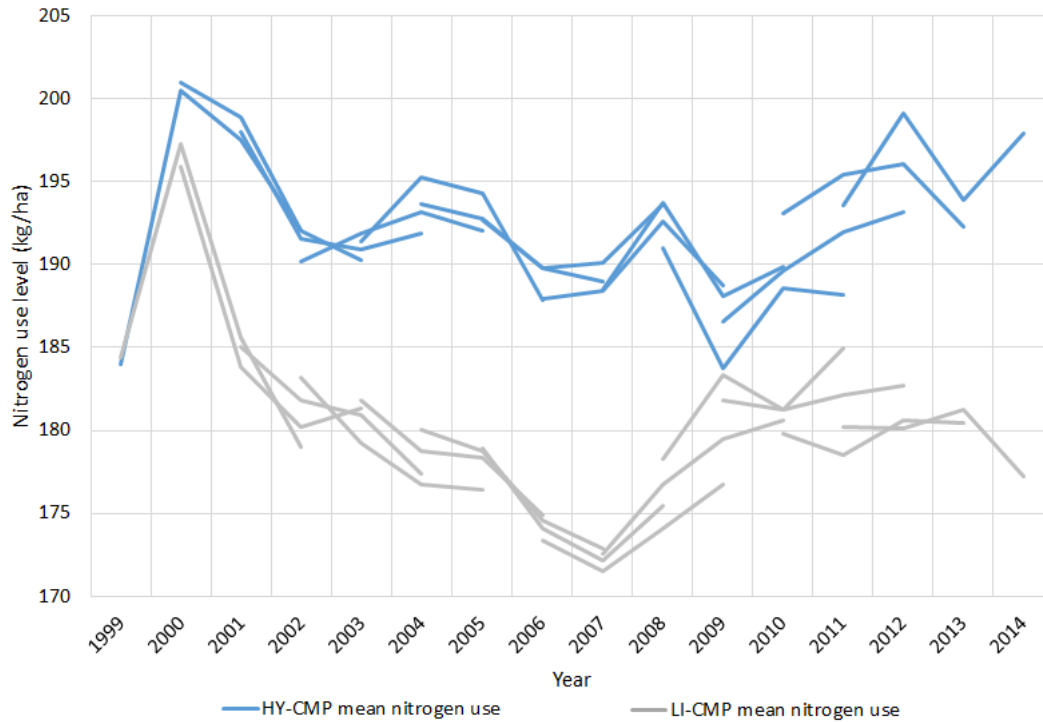
FIGURE 3.21: K-means clustering – Annual mean yield in high-yielding and low-input CMPs



Source: Authors' calculations on CDER data.

FIGURE 3.22: K-means clustering – Annual mean input uses in high-yielding and low-input CMPs

(A) Annual mean nitrogen uses in high-yielding and low-input CMPs



Source: Authors' calculations on CDER data.

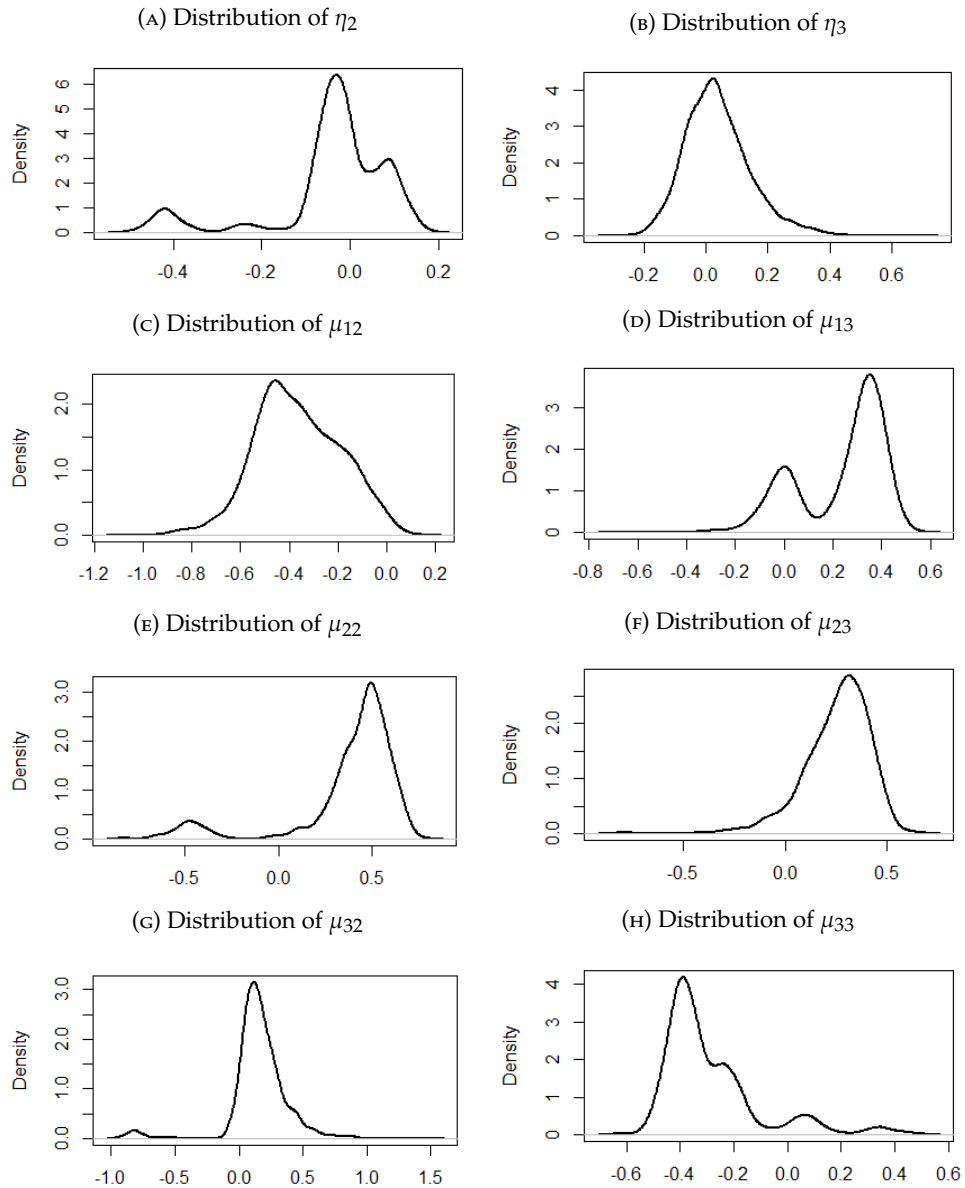
(B) Annual mean pesticide uses in high-yielding and low-input CMPs



Source: Authors' calculations on CDER data.

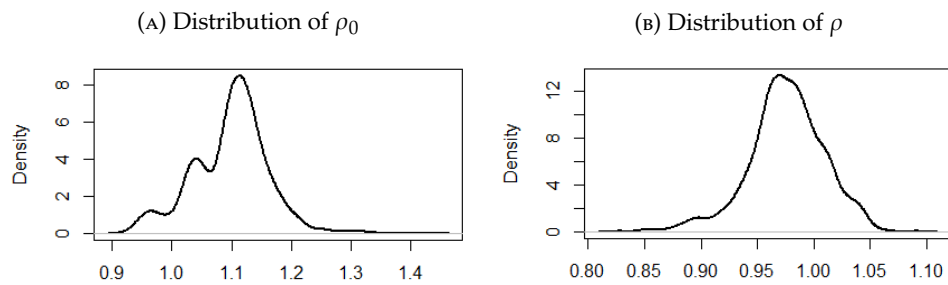
3.9.4 Distribution of random parameters from RPHMM

FIGURE 3.23: Distribution of cost parameters from the technology choice models



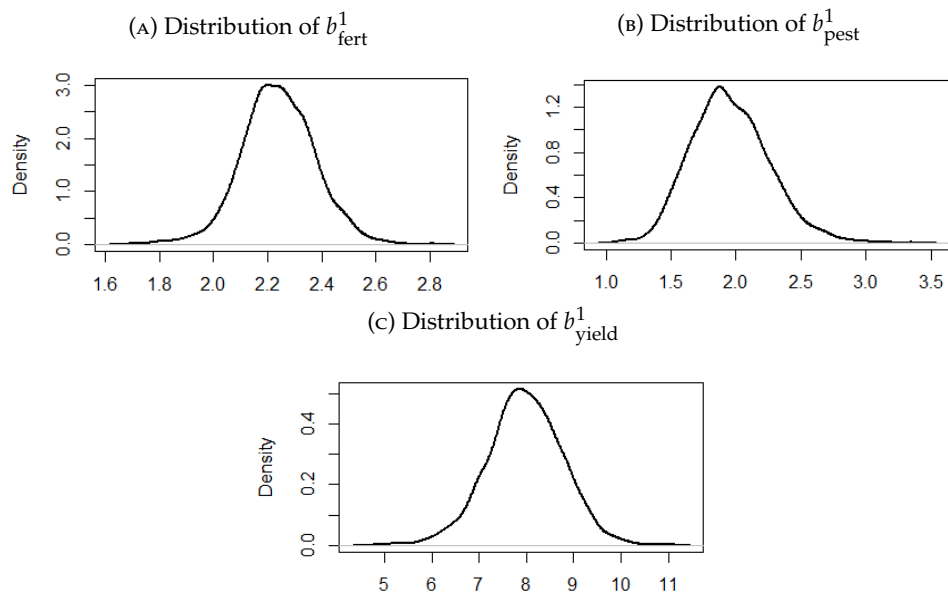
Source: Authors' calculations on CDER data.

FIGURE 3.24: Distribution of scale parameters from the technology choice models



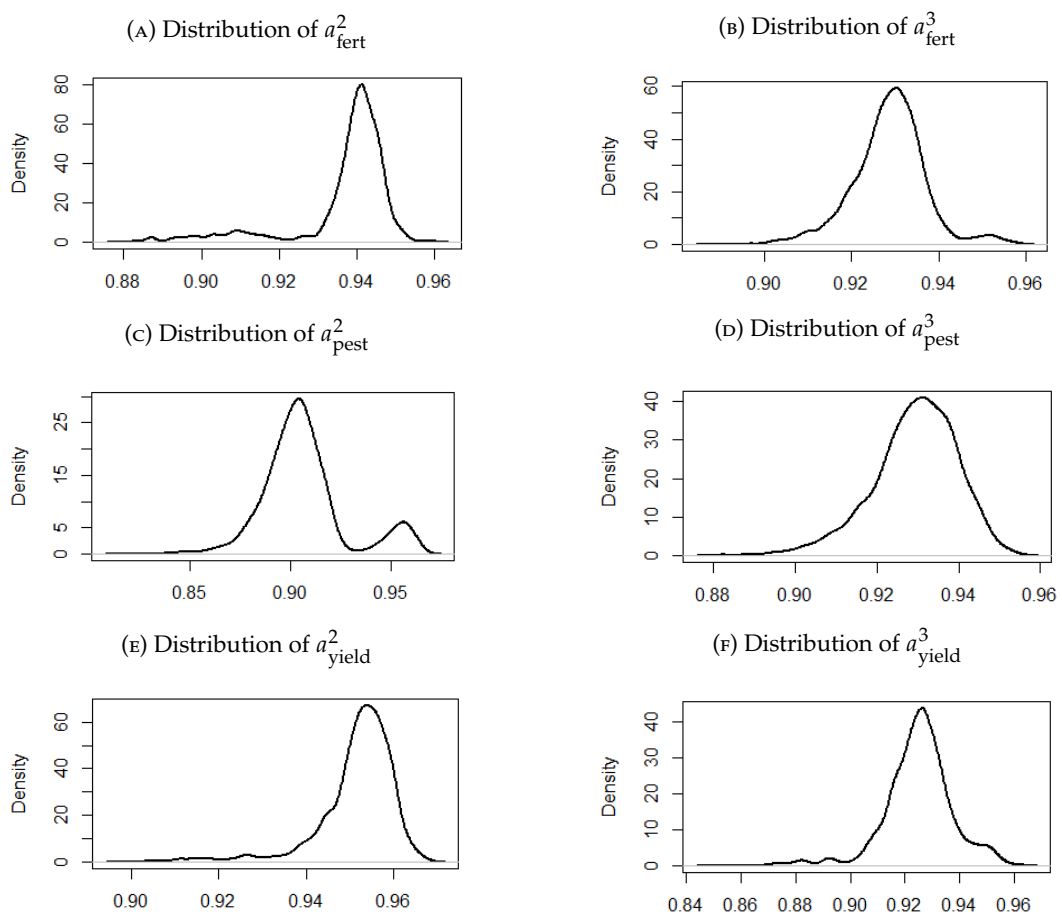
Source: Authors' calculations on CDER data.

FIGURE 3.25: Distribution of the input use and output levels parameters



Source: Authors' calculations on CDER data.

FIGURE 3.26: Distribution of discount parameters from the input use and output level models

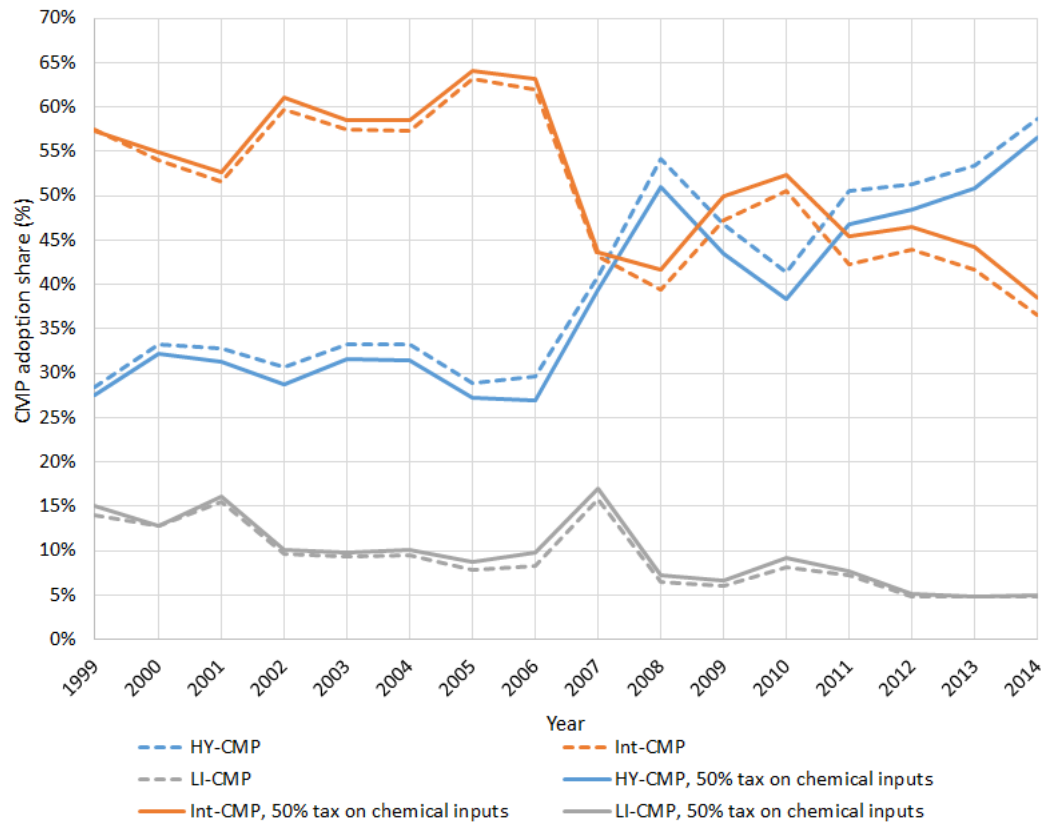


Source: Authors' calculations on CDER data.

3.9.5 Simulation results

As evoked in the core of the article, we performed two types of simulations. First, we simulate a 50% and 100% tax on chemical inputs. We present in Figure 3.27 the results from the 50% tax on chemical inputs.

FIGURE 3.27: Annual change in the CMP adoption share after simulating a 50% tax on chemical inputs

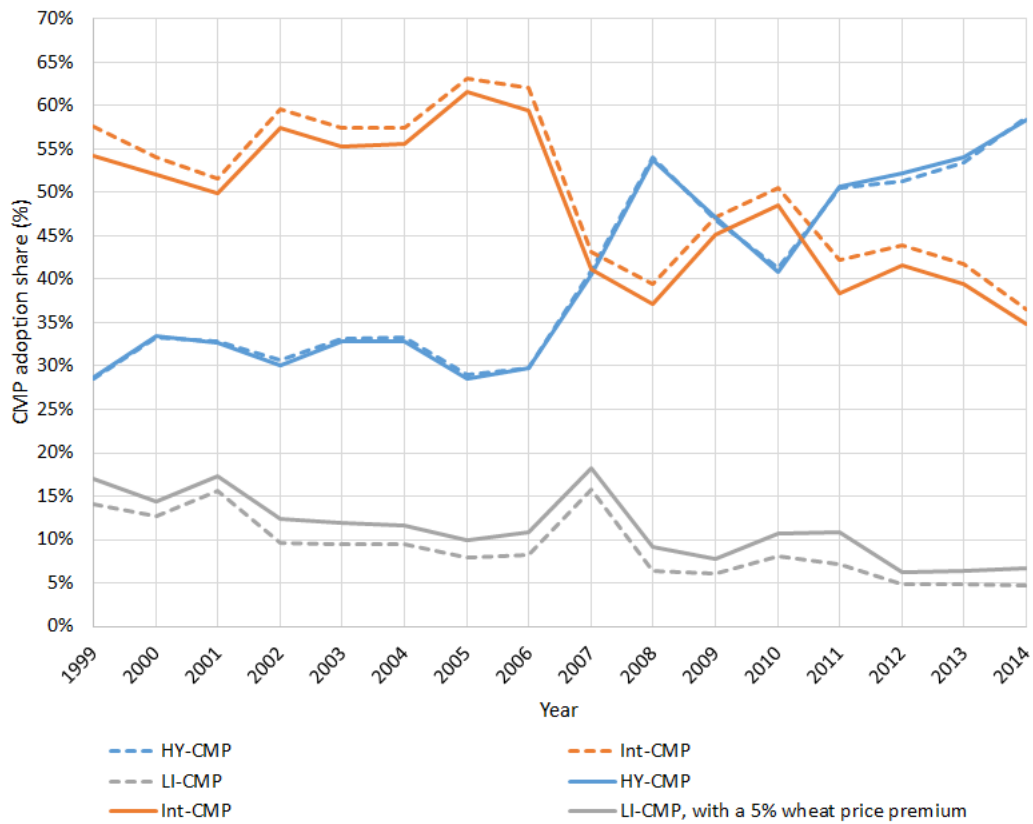


Source: Authors' calculations on CDER data.

Second, we simulated price premiums for low-input wheat farmers. We considered a 5%, 10% and 20% price premium for low-input farmers. Results from the 5% and 20% simulations are presented in Figure 3.28.

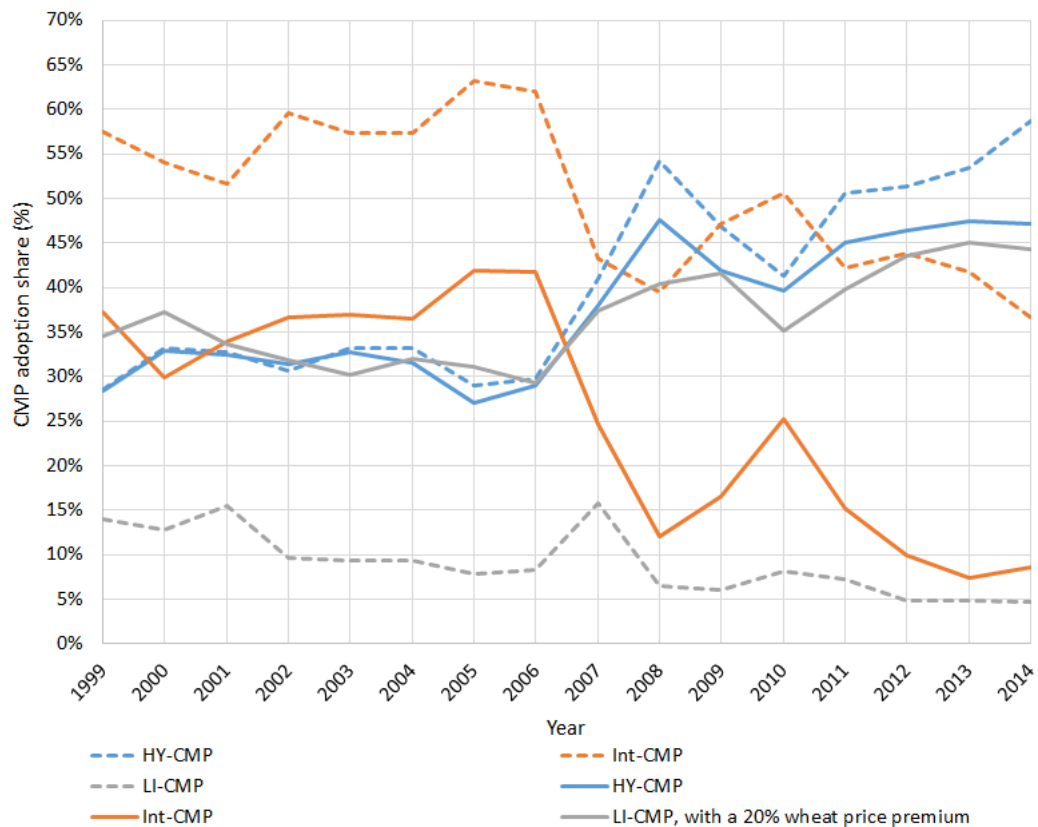
FIGURE 3.28: Annual change in the CMP adoption share after simulating price premiums for low-input wheat producers

(A) Results from the 5% price premium



Source: Authors' calculations on CDER data.

(B) Results from the 20% price premium



Source: Authors' calculations on CDER data.

3.9.6 Determinants of input uses and yield

Regressions were performed in order to see to what extent relative price of inputs (compared to output price, *i.e.* wheat price) are determining the mean yield and mean input use levels obtained with the RPHMM. Results from those regressions are gathered in Table 3.4 and tend to show no significant effect of relative prices on mean yield or input use levels.

TABLE 3.4: Relative prices as non-determining factors of input and yield mean levels

	<i>Yield</i>	<i>Nitrogen</i>	<i>Pesticides</i>
Fertilizer relative price	56.59	-5.44	-8.21
Pesticide relative price	61.59	12.04	20.09
Constante	0.08	0.15	-0.11
Observations	15	15	15
R ²	0.12	0.14	0.08
Residual Std. Error	0.56 (df = 12)	0.06 (df = 12)	0.13 (df = 12)
F Statistic	0.83 (df = 2; 12)	0.95 (df = 2; 12)	0.53 (df = 2; 12)

Note:

*p<0.1; **p<0.05; ***p<0.01

Source: Authors' calculations on CDER data.

General Conclusion

Assessing the impact of technology choice on agricultural production choices or outcomes is a long tradition in the literature. In particular, we chose to assess the impact of two types of cropping management practices (CMPs): a low-input and a high-yielding (or equivalently a high-input) one. Low-input practices were developed in response to the multifaceted noxiousness of pesticides. They rely on agronomic principles that tend to decrease the crop dependence upon chemical inputs. Because they rely on different agronomic principles, we might think that the low-input yield does not respond in the same way to input uses variations than high-input yield levels. Thus, cropping management practices (CMPs) are not only associated to a shift in netputs (*e.g.*, yield and input uses) but rather impact all the parameters in the crop production functions. This encourages us to estimate CMP specific production functions. At the same time, we investigate the determinants of the CMP choice.

The first research question we tried to answer in this PhD thesis is how yield respond to input uses, especially pesticide uses. To answer such question, we considered an extended endogenous regime switching framework for endogenous covariates to estimate CMP specific production functions. In particular, in Chapter 1, we came with an easy to implement, multi-step estimation procedure relying on control functions. This estimation procedure is of particular interest when the modelling framework entails non-linearity issues, as it is the case in Chapter 2. Effectively, in a context where we are particularly interested in pesticides – pesticides being the main target of public authorities – damage abatement functions allow to account for their protective (*vs.* productive) role on yields. Yet, common damage abatement functions entail non-linearity estimation issues. Such non-linearities issues combined to input endogeneity and technology selection issues make the common estimating approach challenging to implement. Hence the proposed estimation approach by Chapter 1 that can be considered as an extension of the well-known two-step approach proposed by Heckman (1976, 1979).

Yet, this estimation procedure deeply relies on the quality of the instrumental variables we use to control for the selection biases and the input use demand models. Unfortunately, our empirical application on Swiss wheat data in Chapter 2 suffers from a weak instrument problem that endanger the control of input endogeneity issues. Thus, estimation results for the production functions are not very reliable.

Results have to be considered cautiously, in particular one should not conclude in terms of how yield respond to input uses variations. Yet, some interesting general results arise from our endogenous regime switching framework. In particular, selection biases seem to affect both the input use and yield models. High-input farmers seem to have larger pesticides uses at baseline because of some specific unobserved characteristics. On the other hand, unobserved specific characteristics of low-input farmers permit them to have higher yield levels. The presence of such selection biases is robust to the different specifications we considered. From a public policy perspective it implies that, to evaluate adequately the impact of the Swiss *Extenso* policy on the reduction of pesticide use and the associated yield loss, this selection bias issue has to be considered.

The second challenge regarding the estimation of CMP specific production functions is how to account for CMPs when they are not observed by economists. Agronomic practices are scarcely available in the data economists usually have at their disposal (*e.g.*, cost accounting data). In Chapter 3, we develop a CMP identification strategy based on yield and input use levels. We also take advantage of the technology adoption dynamics to model technology choice as a Markov process. In particular, we use an hidden Markov model with endogenous switching to identify three CMPs among the winter wheat producers of *La Marne* area. Not only the characteristics of the obtained CMP types are coherent with what is described by agronomists, but also our model informs us about the dynamics of CMP adoption. As expected, the profitability of CMPs heavily depends on wheat prices: the share of profit that is lost to variable input cost is quite insignificant in particular in a context of high crop prices. Indeed, expenses are representing hundred euros per hectare whereas revenues stand at thousand euros. Hence the limited – yet simulated – impact of a pesticide tax on the adoption of low-input practices. Such finding is quite in line with the literature arguing for a low price elasticity of pesticides. From that perspective, to significantly affect farmers' profit, authorities should rely on the revenue part. Indeed, when introducing a price premium for low-input producers, simulations show a greater share of farmers switching to this type of production practices. This revenue strategy was the one adopted by Switzerland with a price premium and a 400 CHF/ha direct payment for *Extenso* wheat producers who make no use of insecticides, fungicides or growth regulators. This strategy might be considered as being quite successful as low-input practices are widely diffused among wheat producers (more than 50% of adoption). From Rogers (1962) innovation adoption theory, one could consider that Swiss wheat farmers that are still using conventional production practices are "laggards" whereas as conventional production practices are still the rule among French wheat producers. Such difference in the adoption rate of low-input practices between France and Switzerland might come from the different public policy strategy they adopted: whereas France chose a pesticide tax scheme

and a voluntary program in the DEPHY network to share information, Switzerland chose to incentivize low-input practices by acting directly on the revenue of wheat producers.

Another interesting result we can derive from Chapter 3 is the role of non-economical drivers in technology adoption. Indeed, even if low-input practices are not interesting from an economic point of view because they are systematically associated to lower profit levels compared to other CMPs, there is still an irreducible small share of farmers choosing them. The importance of non-economical considerations in technology adoption has already been emphasized in the literature by Frey and Stutzer (2006) – for societal/environmental considerations – and Mzoughi (2011) – for social considerations – for instance. Non-economic considerations being great adoption drivers, Frey, Neckermann, et al. (2009) derive the fact that rather than monetary compensation, one should consider award as an instrument to encourage agri-environmental practices.

More generally, this limited difference between low-input and high-input practices in terms of input use questions us about the appeal of low-input production practices, at least for wheat producers. From a benefit-cost point of view, the benefit of low-input practices – *i.e.* a decrease in pesticide use – might appear as being rather limited in comparison to the economical cost it implies for public authorities. The benefit-cost ratio of organic farming might be more interesting in that regard as the decrease in input uses is greater and price premiums for organic farming already exist. In the current economic context, low-input practices are not a viable in-between for wheat producers. We might consider to evaluate the relative benefit-cost ratios of low-input practices *versus* organic practices for wheat producers as well as other crops. Indeed, if the low-input cost-benefit ratio appears to be limited for wheat producers, we might think that for other crops, pesticide use savings are larger and thus might compensate for the revenue loss. In that way, the yield loss might be compensated by public authorities at a lower cost and low-input practices be a viable in-between to conventional or organic production practices.

Bibliography

- Abdulai, A. and W. Huffman (2014). "The adoption and impact of soil and water conservation technology: An endogenous switching regression application". In: *Land economics* 90.1, pp. 26–43.
- Ackerberg, D. et al. (2007). "Econometric tools for analyzing market outcomes". In: *Handbook of econometrics* 6, pp. 4171–4276.
- Ackerberg, D. A., K. Caves, and G. Frazer (2015). "Identification properties of recent production function estimators". In: *Econometrica* 83.6, pp. 2411–2451.
- Acs, S. et al. (2007). "Comparison of conventional and organic arable farming systems in the Netherlands by means of bio-economic modelling". In: *Biological Agriculture & Horticulture* 24.4, pp. 341–361.
- Alene, A. D. and V. M. Manyong (2007). "The effects of education on agricultural productivity under traditional and improved technology in northern Nigeria: an endogenous switching regression analysis". In: *Empirical economics* 32.1, pp. 141–159.
- Altman, R. M. (2007). "Mixed hidden Markov models: an extension of the hidden Markov model to the longitudinal data setting". In: *Journal of the American Statistical Association* 102.477, pp. 201–210.
- Alvarez, A. and J. del Corral (2010). "Identifying different technologies using a latent class model: extensive versus intensive dairy farms". In: *European Review of Agricultural Economics* 37.2, pp. 231–250.
- Amsler, C., A. Prokhorov, and P. Schmidt (2016). "Endogeneity in stochastic frontier models". In: *Journal of Econometrics* 190.2, pp. 280–288.
- Anderson, J. B., D. A. Jolly, and R. D. Green (2005). *Determinants of farmer adoption of organic production methods in the fresh-market produce sector in California: A logistic regression analysis*. Tech. rep.
- Angrist, J. D. and J.-S. Pischke (2008). *Mostly harmless econometrics: An empiricist's companion*. Princeton university press.
- Appleby, A., P. Olson, and D. Colbert (1976). "Winter wheat yield reduction from interference by Italian ryegrass". In: *Agronomy Journal* 68.3, pp. 463–466.
- Asfaw, S. et al. (Sept. 2012). "Poverty Reduction Effects of Agricultural Technology Adoption: A Micro-evidence from Rural Tanzania". In: *Journal of Development Studies - J DEVELOP STUD* 48, pp. 1–18.
- Aubertot, J.-N. et al. (2005). "Pesticides, agriculture et environnement. Réduire l'utilisation des pesticides et en limiter les impacts environnementaux". In:

- Aubertot, J.-N. et al. (2007). *Pesticides, agriculture et environnement. Réduire l'utilisation des pesticides et en limiter les impacts environnementaux. Expertise scientifique collective Inra-Cemagref (décembre 2005)*. Expertises Collectives, 120 p.
- Aune, J. B. (2012). "Conventional, organic and conservation agriculture: production and environmental impact". In: *Agroecology and strategies for climate change*. Springer, pp. 149–165.
- Baan, L. de, S. Spycher, and O. Daniel (2015). "Einsatz von Pflanzenschutzmitteln in der Schweiz von 2009 bis 2012". In: *Agrarforschung Schweiz* 6.2, pp. 48–55.
- Babcock, B. A. (2015). "Using Cumulative Prospect Theory to Explain Anomalous Crop Insurance Coverage Choice". In: *American Journal of Agricultural Economics* 97.5, pp. 1371–1384.
- Badgley, C. et al. (2007). "Organic agriculture and the global food supply". In: *Renewable agriculture and food systems*, pp. 86–108.
- Baffes, J. (2007). *Oil spills on other commodities*. The World Bank.
- Bartholomew, D. J., M. Knott, and I. Moustaki (2011). *Latent variable models and factor analysis: A unified approach*. Vol. 904. John Wiley & Sons.
- Bellemare, M. F., C. B. Barrett, and D. R. Just (2013). "The welfare impacts of commodity price volatility: evidence from rural Ethiopia". In: *American Journal of Agricultural Economics* 95.4, pp. 877–899.
- Bellon, S., F. Lescourret, and J.-P. Calmet (2001). "Characterisation of apple orchard management systems in a French Mediterranean Vulnerable Zone". In: *Agronomie* 21.3, pp. 203–213.
- Benoît, M., F. Le Ber, and J.-F. Mari (2001). "Recherche des successions de cultures et de leurs évolutions : analyse par HMM des données Ter-Uti en Lorraine". In: *Agreste - les cahiers* 31, pp. 23–30.
- Bernard, T. et al. (2016). "Product Market Reforms and Technology Adoption by Senegalese Onion Producers". In:
- Bertrand, M. and T. Doré (2008). "Comment intégrer la maîtrise de la flore adventice dans le cadre général d'un système de production intégrée?" In: *Innovations Agronomiques* 3, pp. 1–13.
- Blazy, J.-M. et al. (2009). "A methodological framework that accounts for farm diversity in the prototyping of crop management systems. Application to banana-based systems in Guadeloupe". In: *Agricultural Systems* 101, pp. 30–41.
- Böcker, T. and R. Finger (2016). "European pesticide tax schemes in comparison: an analysis of experiences and developments". In: *Sustainability* 8.4, p. 378.
- Böcker, T., N. Möhring, and R. Finger (2019). "Herbicide free agriculture? A bio-economic modelling application to Swiss wheat production". In: *Agricultural Systems* 173, pp. 378–392.
- Böcker, T. G. and R. Finger (2017). "A meta-analysis on the elasticity of demand for pesticides". In: *Journal of Agricultural Economics* 68.2, pp. 518–533.
- Boller, E. et al. (1998). "Integrated Production in Europe: 20 years after the declaration of Ovronnaz". In: *IOBC/wprs Bulletin* 21.1.

- Bonomo, M. and R. Garcia (1996). "Consumption and equilibrium asset pricing: An empirical assessment". In: *Journal of Empirical Finance* 3.3, pp. 239–265.
- Booth, J. G. and J. P. Hobert (1999). "Maximizing generalized linear mixed model likelihoods with an automated Monte Carlo EM algorithm". In: *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 61.1, pp. 265–285.
- Booth, J. G., J. P. Hobert, and W. Jank (2001). "A survey of Monte Carlo algorithms for maximizing the likelihood of a two-stage hierarchical model". In: *Statistical Modelling* 1.4, pp. 333–349.
- Boquet, D. J. and C. C. Johnson (1987). "Fertilizer Effects on Yield, Grain Composition, and Foliar Disease of Doublecrop Soft Red Winter Wheat 1". In: *Agronomy journal* 79.1, pp. 135–141.
- Bostian, A. et al. (2019). "Assessing the productivity consequences of agri-environmental practices when adoption is endogenous". In: *Journal of Productivity Analysis*, pp. 1–22.
- Bouchard, C. et al. (2008). "Associer des itinéraires techniques de niveau d'intrants variés à des variétés rustiques de blé tendre: évaluation économique, environnementale et énergétique". In: *Le Courrier de l'Environnement de l'INRA* 55.55, pp. 53–77.
- Bravo-Ureta, B. E., W. Greene, and D. Solís (2012). "Technical efficiency analysis correcting for biases from observed and unobserved variables: an application to a natural resource management project". In: *Empirical Economics* 43.1, pp. 55–72.
- Bravo-Ureta, B. E. et al. (2021). "Technology and Technical Efficiency Change: Evidence from a Difference in Differences Selectivity Corrected Stochastic Production Frontier Model". In: *American Journal of Agricultural Economics* 103.1, pp. 362–385.
- Brisson, N. et al. (2010). "Why are wheat yields stagnating in Europe? A comprehensive data analysis for France". In: *Field Crops Research* 119.1, pp. 201–212.
- Burrows, T. M. (1983). "Pesticide demand and integrated pest management: a limited dependent variable analysis". In: *American Journal of Agricultural Economics* 65.4, pp. 806–810.
- Burton, M., D. Rigby, and T. Young (1999). "Analysis of the Determinants of Adoption of Organic Horticultural Techniques in the UK". In: *Journal of Agricultural Economics* 50.1, pp. 47–63.
- Butault, J.-P. et al. (2011). "L'utilisation des pesticides en France: état des lieux et perspectives de réduction". In:
- Cappé, O., E. Moulines, and T. Rydén (2006). *Inference in hidden Markov models*. Springer Science & Business Media.
- Carpentier, A. and E. Letort (2011). "Accounting for heterogeneity in multicrop microeconomic models: implications for variable input demand modeling". In: *American journal of agricultural economics* 94.1, pp. 209–224.
- (2014). "Multicrop production models with Multinomial Logit acreage shares". In: *Environmental and Resource Economics* 59.4, pp. 537–559.

- Carpentier, A. and R. D. Weaver (1997). "Damage control productivity: Why econometrics matters". In: *American Journal of Agricultural Economics* 79.1, pp. 47–61.
- Carson, R. (2002). *Silent spring*. Houghton Mifflin Harcourt.
- Cecchetti, S. G., P.-s. Lam, and N. Mark (1988). *Mean reversion in equilibrium asset prices*.
- Chabé-Ferret, S. and J. Subervie (2013). "How much green for the buck? Estimating additional and windfall effects of French agro-environmental schemes by DID-matching". In: *Journal of Environmental Economics and Management* 65.1, pp. 12–27.
- Chamberlain, G. (1982). "Multivariate regression models for panel data". In: *Journal of econometrics* 18.1, pp. 5–46.
- (1984). "Panel data". In: *Handbook of econometrics* 2, pp. 1247–1318.
- (1987). "Asymptotic efficiency in estimation with conditional moment restrictions". In: *Journal of econometrics* 34.3, pp. 305–334.
- Chambers, R. G. et al. (1988). *Applied production analysis: a dual approach*. Cambridge University Press.
- Chauvet, M. and J. D. Hamilton (2006). "Chapter 1 Dating Business Cycle Turning Points". In: *Nonlinear Time Series Analysis of Business Cycles*. Ed. by C. Milas, P. Rothman, and D. van Dijk. Vol. 276. Contributions to Economic Analysis. Elsevier, pp. 1–54.
- Chavas, J.-P. and M. T. Holt (1990). "Acreage decisions under risk: the case of corn and soybeans". In: *American Journal of Agricultural Economics* 72.3, pp. 529–538.
- Chavas, J.-P. and C. Nauges (2020). "Uncertainty, learning, and technology adoption in agriculture". In: *Applied Economic Perspectives and Policy* 42.1, pp. 42–53.
- Chen, P.-Y. et al. (2012). "Modelling the Effects of Oil Prices on Global Fertilizer Prices and Volatility". In: *Journal of Risk and Financial Management* 5.1, pp. 78–114.
- Clark, M. S. et al. (1998). "Changes in soil chemical properties resulting from organic and low-input farming practices". In: *Agronomy Journal* 90.5, pp. 662–671.
- Clark, S. et al. (1999). "Crop-yield and economic comparisons of organic, low-input, and conventional farming systems in California's Sacramento Valley". In: *American journal of alternative agriculture*, pp. 109–121.
- Colla, G. et al. (2000). "Soil physical properties and tomato yield and quality in alternative cropping systems". In: *Agronomy Journal* 92.5, pp. 924–932.
- Connor, D. (2008). "Organic agriculture cannot feed the world". In: *Field Crops Research* 106.2, p. 187.
- Cowan, R. and P. Gunby (1996). "Sprayed to death: path dependence, lock-in and pest control strategies". In: *The economic journal* 106.436, pp. 521–542.
- Dakpo, K. H. et al. (2021). "Latent Class Modelling for a Robust Assessment of Productivity: Application to French Grazing Livestock Farms". In: *Journal of Agricultural Economics*.
- De Janvry, A., K. Macours, and E. Sadoulet (2017). *Learning for adopting: Technology adoption in developing country agriculture*. FERDI.

- Delattre, M. and M. Lavielle (2012). "Maximum likelihood estimation in discrete mixed hidden Markov models using the SAEM algorithm". In: *Computational Statistics & Data Analysis* 56.6, pp. 2073–2085.
- Delyon, B., M. Lavielle, and E. Moulines (1999). "Convergence of a Stochastic Approximation Version of the EM Algorithm". In: *The Annals of Statistics* 27.1, pp. 94–128.
- Dempster, A. P., N. M. Laird, and D. B. Rubin (1977). "Maximum Likelihood from Incomplete Data Via the EM Algorithm". In: *Journal of the Royal Statistical Society: Series B (Methodological)* 39.1, pp. 1–22.
- Di Falco, S. and M. Veronesi (2013). "How can African agriculture adapt to climate change? A counterfactual analysis from Ethiopia". In: *Land Economics* 89.4, pp. 743–766.
- Di Falco, S., M. Veronesi, and M. Yesuf (2011). "Does adaptation to climate change provide food security? A micro-perspective from Ethiopia". In: *American Journal of Agricultural Economics* 93.3, pp. 829–846.
- Directive 2009/128/EC of the European Parliament and of the Council of 21 October 2009 establishing a framework for Community action to achieve the sustainable use of pesticides (2009). Official Journal of the European Union.
- Efron, B. and R. Tibshirani (1986). "Bootstrap methods for standard errors, confidence intervals, and other measures of statistical accuracy". In: *Statistical science*, pp. 54–75.
- El Benni, N. and R. Finger (2013). "The effect of agricultural policy reforms on income inequality in Swiss agriculture-An analysis for valley, hill and mountain regions". In: *Journal of Policy Modeling* 35.4, pp. 638–651.
- Feder, G. and D. L. Umali (1993). "The adoption of agricultural innovations: a review". In: *Technological forecasting and social change* 43.3-4, pp. 215–239.
- Féménia, F. and E. Letort (2016). "How to significantly reduce pesticide use: An empirical evaluation of the impacts of pesticide taxation associated with a change in cropping practice". In: *Ecological Economics* 125, pp. 27–37.
- Fernandez-Cornejo, J. (1996). "The microeconomic impact of IPM adoption: theory and application". In: *Agricultural and Resource Economics Review* 25.2, pp. 149–160.
- (1998). "Environmental and economic consequences of technology adoption: IPM in viticulture". In: *Agricultural economics* 18.2, pp. 145–155.
- Fernandez-Cornejo, J., C. Hendricks, and A. Mishra (2005). "Technology adoption and off-farm household income: the case of herbicide-tolerant soybeans". In: *Journal of Agricultural and Applied Economics* 37.3, pp. 549–563.
- Fernandez-Cornejo, J. and S. J. Wechsler (2012). "Revisiting the impact of Bt corn adoption by US farmers". In: *Agricultural and Resource Economics Review* 41, pp. 377–390.
- Fess, T. L., J. B. Kotcon, and V. A. Benedito (2011). "Crop breeding for low input agriculture: a sustainable response to feed a growing world population". In: *Sustainability* 3.10, pp. 1742–1772.

- Finger, R. (2008). "Impacts of agricultural policy reforms on crop yields". In: *Euro-Choices* 7.3, pp. 24–25.
- (2014). "Risk considerations in the economic assessment of low-input crop production techniques: an example from Swiss wheat production". In: *International journal of agricultural resources, governance and ecology* 10.1, pp. 63–77.
- Finger, R. and N. El Benni (2013). "Farmers' adoption of extensive wheat production—Determinants and implications". In: *Land Use Policy* 30.1, pp. 206–213.
- Finger, R. et al. (2017). "Revisiting pesticide taxation schemes". In: *Ecological Economics* 134, pp. 263–266.
- Foley, J. A. et al. (2005). "Global consequences of land use". In: *science* 309.5734, pp. 570–574.
- Foley, J. A. et al. (2011). "Solutions for a cultivated planet". In: *Nature* 478.7369, pp. 337–342.
- Foster, A. D. and M. R. Rosenzweig (2010). "Microeconomics of technology adoption". In: *Annual Revue of Economics* 2.1, pp. 395–424.
- Frey, B. S., S. Neckermann, et al. (2009). "Awards: A disregarded source of motivation". In: *Rationality, Markets and Morals* 11.
- Frey, B. S. and A. Stutzer (2006). "Environmental morale and motivation".
- Frisvold, G. B. (2019). "How low can you go? Estimating impacts of reduced pesticide use". In: *Pest management science* 75.5, pp. 1223–1233.
- Fuglie, K. O. and D. J. Bosch (1995). "Economic and environmental implications of soil nitrogen testing: A switching-regression analysis". In: *American Journal of Agricultural Economics* 77.4, pp. 891–900.
- Gardebroek, C. (Dec. 2006). "Comparing risk attitudes of organic and non-organic farmers with a Bayesian random coefficient model". In: *European Review of Agricultural Economics* 33.4, pp. 485–510.
- Gardebroek, C., M. D. Chavez, and A. O. Lansink (2010). "Analysing Production Technology and Risk in Organic and Conventional Dutch Arable Farming using Panel Data". In: *Journal of Agricultural Economics* 61.1, pp. 60–75.
- Gebauer, R. H. (1987). "Socio-economic classification of farm households —conceptual, methodical and empirical considerations". In: *European Review of Agricultural Economics* 14.3, pp. 261–283.
- Genius, M. et al. (2014). "Information transmission in irrigation technology adoption and diffusion: Social learning, extension services, and spatial effects". In: *American Journal of Agricultural Economics* 96.1, pp. 328–344.
- Goetz, R.-U., H. Schmid, and B. Lehmann (2006). "Determining the economic gains from regulation at the extensive and intensive margins". In: *European review of agricultural economics* 33.1, pp. 1–30.
- Goldfeld, S. M. and R. E. Quandt (1973). "A Markov model for switching regressions". In: *Journal of Econometrics* 1.1, pp. 3–15.
- Goring, C. and D. Laskowski (1982). "The effects of pesticides on nitrogen transformations in soils". In: *Nitrogen in agricultural soils* 22, pp. 689–720.

- Greene, W. (2005). "Reconsidering heterogeneity in panel data estimators of the stochastic frontier model". In: *Journal of econometrics* 126.2, pp. 269–303.
- (2010). "A stochastic frontier model with correction for sample selection". In: *Journal of productivity analysis* 34.1, pp. 15–24.
- Greene, W. H. (2020). *Econometric Analysis*. Pearson Education Limited.
- Greene, W. H. and M. Liu (2021). "Review of difference-in-difference analyses in social sciences: application in policy test research". In: *Handbook of financial econometrics, mathematics, statistics, and machine learning*. World Scientific, pp. 4255–4280.
- Greiner, R., L. Patterson, and O. Miller (2009). "Motivations, risk perceptions and adoption of conservation practices by farmers". In: *Agricultural systems* 99.2-3, pp. 86–104.
- Griliches, Z. (1957). "Hybrid corn: An exploration in the economics of technological change". In: *Econometrica, Journal of the Econometric Society*, pp. 501–522.
- Griliches, Z. and J. Mairesse (1995). "Production functions: the search for identification". In: *NBER working paper w5067*.
- Gu, M. G. and H.-T. Zhu (2001). "Maximum likelihood estimation for spatial models by Markov chain Monte Carlo stochastic approximation". In: *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 63.2, pp. 339–355.
- Guay, E. (2012). "Étude de l'impact sur la flore adventice de plusieurs systèmes de culture à bas niveau d'intrants". MA thesis. Ecole d'Ingénieurs de Purpan (Toulouse, France).
- Haavelmo, T. (1943). "The statistical implications of a system of simultaneous equations". In: *Econometrica, Journal of the Econometric Society*, pp. 1–12.
- Hamilton, J. D. (1989). "A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle". In: *Econometrica* 57.2, pp. 357–384.
- Hartigan, J. A. and M. A. Wong (1979). "A K-means clustering algorithm". In: *Applied Statistics* 28, pp. 100–108.
- Heckman, J., J. L. Tobias, and E. Vytlacil (2003). "Simple estimators for treatment parameters in a latent-variable framework". In: *Review of Economics and Statistics* 85.3, pp. 748–755.
- Heckman, J. J. (1976). "The common structure of statistical models of truncation, sample selection and limited dependent variables and a simple estimator for such models". In: *Annals of economic and social measurement, volume 5, number 4*. NBER, pp. 475–492.
- (1978). "Dummy endogenous variables in a simultaneous equation system". In: *Econometrica: Journal of the Econometric Society*, pp. 931–959.
- (1979). "Sample selection bias as a specification error". In: *Econometrica: Journal of the econometric society*, pp. 153–161.
- Heckman, J. J. and B. E. Honore (1990). "The empirical content of the Roy model". In: *Econometrica: Journal of the Econometric Society*, pp. 1121–1149.

- Henson, J. F. and L. S. Jordan (1982). "Wild oat (*Avena fatua*) competition with wheat (*Triticum aestivum* and *T. turgidum durum*) for nitrate". In: *Weed Science* 30.3, pp. 297–300.
- Hossard, L. et al. (2017). "Lack of evidence for a decrease in synthetic pesticide use on the main arable crops in France". In: *Science of The Total Environment* 575, pp. 152–161.
- Howard, D., A. Chambers, and J. Logan (1994). "Nitrogen and fungicide effects on yield components and disease severity in wheat". In: *Journal of production agriculture* 7.4, pp. 448–454.
- Howley, P. et al. (2015). "Explaining the economic 'irrationality' of farmers' land use behaviour: The role of productivist attitudes and non-pecuniary benefits". In: *Ecological Economics* 109, pp. 186–193.
- Husson, F., J. Josse, and J. Pagès (2010). *Principal component methods - hierarchical clustering - partitional clustering: why would we need to choose for visualizing data?* Tech. rep. Agrocampus Ouest.
- Hwu, S.-T., C.-J. Kim, and J. Piger (2019). "An n-state endogenous markovswitching model with applications in macroeconomics and finance". In: *Manuscript, University of Washington*.
- Imbens, G. W. and J. M. Wooldridge (2009). "Recent developments in the econometrics of program evaluation". In: *Journal of economic literature* 47.1, pp. 5–86.
- Just, R. E., J. M. Antle, et al. (1990). "Interactions between agricultural and environmental policies: a conceptual framework". In: *American Economic Review* 80.2, pp. 197–202.
- Just, R. E. and R. D. Pope (1978). "Stochastic specification of production functions and economic implications". In: *Journal of econometrics* 7.1, pp. 67–86.
- Kabunga, N. S., T. Dubois, and M. Qaim (2012). "Heterogeneous information exposure and technology adoption: the case of tissue culture bananas in Kenya". In: *Agricultural Economics* 43.5, pp. 473–486.
- Kaufman, L. and P. J. Rousseeuw (2009). *Finding groups in data: an introduction to cluster analysis*. Vol. 344. John Wiley & Sons.
- Khanal, U. et al. (2018a). "Climate change adaptation strategies and food productivity in Nepal: a counterfactual analysis". In: *Climatic change* 148.4, pp. 575–590.
- Khanal, U. et al. (2018b). "Farmers' adaptation to climate change, its determinants and impacts on rice yield in Nepal". In: *Ecological Economics* 144, pp. 139–147.
- Khanna, M. (2001). "Sequential adoption of site-specific technologies and its implications for nitrogen productivity: A double selectivity model". In: *American journal of agricultural economics* 83.1, pp. 35–51.
- Kim, C.-J., J. Piger, and R. Startz (2008). "Estimation of Markov regime-switching regression models with endogenous switching". In: *Journal of Econometrics* 143.2, pp. 263–273.
- Klein, L. R. (1953). *A Textbook of Econometrics*. Row, Peterson.

- Kong, A. Y. et al. (2011). "Microbial community composition and carbon cycling within soil microenvironments of conventional, low-input, and organic cropping systems". In: *Soil Biology and Biochemistry* 43.1, pp. 20–30.
- Koutchadé, O. P., A. Carpentier, and F. Féménia (2018). "Modeling heterogeneous farm responses to european union biofuel support with a random parameter multicrop model". In: *American Journal of Agricultural Economics* 100.2, pp. 434–455.
- (2020). *Crop choices in micro-econometric multi-crop models: modelling corners, kinks and jumps*. Tech. rep.
- Kudsk, P., L. N. Jørgensen, and J. E. Ørum (2018). "Pesticide Load—A new Danish pesticide risk indicator with multiple applications". In: *Land Use Policy* 70, pp. 384–393.
- Kuhn, E. and M. Lavielle (2005). "Maximum likelihood estimation in nonlinear mixed effects models". In: *Computational statistics & data analysis* 49.4, pp. 1020–1038.
- Kumbhakar, S. C., C. F. Parmeter, and V. Zelenyuk (2020). "Stochastic frontier analysis: Foundations and advances I". In: *Handbook of production economics*, pp. 1–39.
- Kumbhakar, S. C., E. G. Tsionas, and T. Sipiläinen (2009). "Joint estimation of technology choice and technical efficiency: an application to organic and conventional dairy farming". In: *Journal of Productivity Analysis* 31.3, pp. 151–161.
- Kutlu, L. (2010). "Battese-Coelli estimator with endogenous regressors". In: *Economics Letters* 109.2, pp. 79–81.
- Lansink, A. O. (1999). "Generalised maximum entropy estimation and heterogeneous technologies". In: *European review of agricultural economics* 26.1, pp. 101–115.
- Lansink, A. O., K. Pietola, and S. Bäckman (2002). "Efficiency and productivity of conventional and organic farms in Finland 1994–1997". In: *European Review of Agricultural Economics* 29.1, pp. 51–65.
- Larédo, P. and A. Hocdé (2014). *Variétés rustiques et itinéraires techniques économes en intrants*. Tech. rep. ASIRPA.
- Latruffe, L. and C. Nauges (2014). "Technical efficiency and conversion to organic farming: the case of France". In: *European Review of Agricultural Economics* 41.2, pp. 227–253.
- Latruffe, L. et al. (2017). "Subsidies and technical efficiency in agriculture: Evidence from European dairy farms". In: *American Journal of Agricultural Economics* 99.3, pp. 783–799.
- Lavielle, M. (2018). "Pharmacometrics models with hidden Markovian dynamics". In: *Journal of pharmacokinetics and pharmacodynamics* 45.1, pp. 91–105.
- Lavielle, M. and C. Mbogning (2014). "An improved SAEM algorithm for maximum likelihood estimation in mixtures of non linear mixed effects models". In: *Statistics and Computing* 24.5, pp. 693–707.
- Lechenet, M. et al. (2017). "Reducing pesticide use while preserving crop productivity and profitability on arable farms". In: *Nature Plants* 3.3, pp. 1–6.

- Lee, L.-F. (1978). "Unionism and wage rates: A simultaneous equations model with qualitative and limited dependent variables". In: *International economic review*, pp. 415–433.
- (1982). "Some approaches to the correction of selectivity bias". In: *The Review of Economic Studies* 49.3, pp. 355–372.
- Lemieux, T. (1998). "Estimating the effects of unions on wage inequality in a panel data model with comparative advantage and nonrandom selection". In: *Journal of Labor Economics* 16.2, pp. 261–291.
- Levinsohn, J. and A. Petrin (2003). "Estimating production functions using inputs to control for unobservables". In: *The review of economic studies* 70.2, pp. 317–341.
- Lichtenberg, E. (2004). "Some hard truths about agriculture and the environment". In: *Agricultural and Resource Economics Review* 33.1, pp. 24–33.
- Lichtenberg, E. and D. Zilberman (1986). "The econometrics of damage control: why specification matters". In: *American Journal of Agricultural Economics* 68.2, pp. 261–273.
- Liebhardt, W. et al. (1989). "Crop production during conversion from conventional to low-input methods". In: *Agronomy Journal* 81.2, pp. 150–159.
- Lin, B. B. (2011). "Resilience in agriculture through crop diversification: adaptive management for environmental change". In: *BioScience* 61.3, pp. 183–193.
- Lintell Smith, G et al. (1992). *The effects of reduced nitrogen and weed competition on the yield of winter wheat*. Tech. rep.
- Livingston, M., J. Fernandez-Cornejo, and G. B. Frisvold (2016). "Economic returns to herbicide resistance management in the short and long run: the role of neighbor effects". In: *Weed Science* 64.sp1, pp. 595–608.
- Livingston, M. et al. (2015). "The economics of glyphosate resistance management in corn and soybean production (ERR-184). Washington, DC: US Department of Agriculture". In: *Economic Research Service*.
- Los, E., C. Gardebroek, and R. Huirne (2021). "Firm-specific responses to energy policies in Dutch horticulture". In: *European Review of Agricultural Economics* 48.2, pp. 362–384.
- Loyce, C. and J. Meynard (1997). "Low input wheat management techniques are more efficient in ethanol production". In: *Industrial Crops and Products* 6.3-4, pp. 271–283.
- Loyce, C. et al. (2008). "Interaction between cultivar and crop management effects on winter wheat diseases, lodging, and yield". In: *Crop protection* 27.7, pp. 1131–1142.
- Loyce, C. et al. (2012). "Growing winter wheat cultivars under different management intensities in France: A multicriteria assessment based on economic, energetic and environmental indicators". In: *Field crops research* 125, pp. 167–178.
- Lucas, P. (2007). "Le concept de la protection intégrée des cultures". In: *Innovations agronomiques* 1, pp. 15–21.

- MacQueen, J. et al. (1967). "Some methods for classification and analysis of multivariate observations". In: *Proceedings of the fifth Berkeley symposium on mathematical statistics and probability*. Vol. 1. 14. Oakland, CA, USA, pp. 281–297.
- Mäder, P. et al. (2000). "Arbuscular mycorrhizae in a long-term field trial comparing low-input (organic, biological) and high-input (conventional) farming systems in a crop rotation". In: *Biology and fertility of Soils* 31.2, pp. 150–156.
- Mahé, L.-P. and P. Rainelli (1987). "Impact des pratiques et des politiques agricoles sur l'environnement". In: *Cahiers d'Economie et de Sociologie Rurales (CESR)* 4.905-2016-70149, p. 9.
- Marra, M. and A. A. Pannell David J .and Ghadim (2003). "The economics of risk, uncertainty and learning in the adoption of new agricultural technologies: where are we on the learning curve?" In: *Agricultural systems* 75.2-3, pp. 215–234.
- Marschak, J. and W. H. Andrews (1944). "Random simultaneous equations and the theory of production". In: *Econometrica, Journal of the Econometric Society*, pp. 143–205.
- Martinez Cillero, M. et al. (2018). "Technology heterogeneity and policy change in farm-level efficiency analysis: an application to the Irish beef sector". In: *European Review of Agricultural Economics* 46.2, pp. 193–214.
- Maruotti, A. (2011). "Mixed hidden markov models for longitudinal data: An overview". In: *International Statistical Review* 79.3, pp. 427–454.
- Maseda, F., F. Diaz, and C. Alvarez (2004). "Family Dairy Farms in Galicia (N. W. Spain): Classification by Some Family and Farm Factors Relevant to Quality of Life". In: *Biosystems Engineering* 87.4, pp. 509–521.
- Matson, P. A. et al. (1997). "Agricultural intensification and ecosystem properties". In: *Science* 277.5325, pp. 504–509.
- Matyjaszczyk, E. (2019). "Problems of implementing compulsory integrated pest management". In: *Pest management science* 75.8, pp. 2063–2067.
- Mayen, C. D., J. V. Balagtas, and C. E. Alexander (2010). "Technology adoption and technical efficiency: organic and conventional dairy farms in the United States". In: *American Journal of Agricultural Economics* 92.1, pp. 181–195.
- McLachlan, G. and T. Krishnan (2007). *The EM algorithm and extensions*. Vol. 382. John Wiley & Sons.
- McLachlan, G. and D. Peel (2004). *Finite mixture models*. John Wiley & Sons.
- Mennig, P. and J. Sauer (2019). "The impact of agri-environment schemes on farm productivity: a DID-matching approach". In: *European Review of Agricultural Economics* 47.3, pp. 1045–1093.
- (2020). "The impact of agri-environment schemes on farm productivity: a DID-matching approach". In: *European Review of Agricultural Economics* 47.3, pp. 1045–1093.
- Meynard, J.-M. (1985). "Construction d'itinéraires techniques pour la conduite du blé d'hiver". PhD thesis.
- (2008). *Introduction générale: Produire autrement: réinventer les systèmes de culture*.

- Meynard, J.-M. and P. Girardin (1991). "Produire autrement". In: *Courrier de la Cellule Environnement INRA* 15.15, pp. 1–19.
- Meynard, J. (1991). "Pesticides et itinéraires techniques". In: *Phytoprotection des plantes biopesticides*. Paris: INRA, pp. 85–100.
- Michler, J. et al. (2018). "Money matters: The role of yields and profits in agricultural technology adoption". In:
- Michler, J. D. and G. E. Shively (2015). "Land Tenure, Tenure Security and Farm Efficiency: Panel Evidence from the Philippines". In: *Journal of Agricultural Economics* 66.1, pp. 155–169.
- Miller, N. et al. (July 2017). "Farmers' Adoption Path of Precision Agriculture Technology". In: *Advances in Animal Biosciences* 8, pp. 708–712.
- Mischler, P. et al. (2009). "Huit fermes de grande culture engagées en production intégrée réduisent les pesticides sans baisse de marge". In: *Le Courrier de l'environnement de l'INRA* 57.57, pp. 73–91.
- Möhring, N., S. Gaba, and R. Finger (2019). "Quantity based indicators fail to identify extreme pesticide risks". In: *Science of the Total Environment* 646, pp. 503–523.
- Möhring, N. et al. (2020). "Are pesticides risk decreasing? The relevance of pesticide indicator choice in empirical analysis". In: *Agricultural Economics* 51.3, pp. 429–444.
- Monolix Methodology* (2014). Version 4.3.3.
- Mouron, P. and D. Schmid (2012). "Grundlagenbericht 2011". In: *Agroscope, Ettenhausen*.
- Mundlak, Y. (1961). "Empirical production function free of management bias". In: *Journal of Farm Economics* 43.1, pp. 44–56.
- (1978). "On the pooling of time series and cross section data". In: *Econometrica: journal of the Econometric Society*, pp. 69–85.
- (1996). "Production function estimation: reviving the primal". In: *Econometrica: Journal of the Econometric Society*, pp. 431–438.
- (2001). "Chapter 1 Production and supply". In: *Agricultural Production*. Vol. 1. Handbook of Agricultural Economics. Elsevier, pp. 3–85.
- Mzoughi, N. (2011). "Farmers adoption of integrated crop protection and organic farming: Do moral and social concerns matter?" In: *Ecological Economics* 70.8, pp. 1536–1545.
- Newey, W. K. (1990). "Efficient instrumental variables estimation of nonlinear models". In: *Econometrica: Journal of the Econometric Society*, pp. 809–837.
- (1993). "16 Efficient estimation of models with conditional moment restrictions". In: *Econometrics*. Vol. 11. Handbook of Statistics. Elsevier, pp. 419–454.
- Olley, G. S. and A. Pakes (1996). "The dynamics of productivity in the telecommunications equipment industry". In: *Econometrica* 64.6, p. 1263.
- Orea, L. and S. C. Kumbhakar (2004). "Efficiency measurement using a latent class stochastic frontier model". In: *Empirical economics* 29.1, pp. 169–183.

- Orea, L., J. A. Perez, and D. Roibas (2015). "Evaluating the double effect of land fragmentation on technology choice and dairy farm productivity: A latent class model approach". In: *Land Use Policy* 45, pp. 189–198.
- Pedersen, A. B. et al. (2012). "Optimising the effect of policy instruments: a study of farmers' decision rationales and how they match the incentives in Danish pesticide policy". In: *Journal of environmental planning and management* 55.8, pp. 1094–1110.
- Pell, M., B. Stenberg, and L. Torstensson (1998). "Potential denitrification and nitrification tests for evaluation of pesticide effects in soil". In: *Ambio*, pp. 24–28.
- Perrot, C. (1990). "Typologie d'exploitations construite par agrégation autour de pôles définis à dire d'experts : Proposition méthodologique et premiers résultats obtenus en Haute-Marne". In: *INRA Productions animales* 3.1, pp. 51–66.
- Pietola, K. S. and A. O. Lansink (2001). "Farmer response to policies promoting organic farming technologies in Finland". In: *European Review of Agricultural Economics* 28.1, pp. 1–15.
- Pimentel, D. and A. Greiner (1997). "Environmental and socio-economic costs of pesticide use". In: *Techniques for reducing pesticide use: Economic and environmental benefits*, pp. 51–78.
- Pimentel, D. et al. (1992). "Environmental and economic costs of pesticide use". In: *BioScience* 42.10, pp. 750–760.
- Piñeiro, V. et al. (2020). "A scoping review on incentives for adoption of sustainable agricultural practices and their outcomes". In: *Nature Sustainability* 3.10, pp. 809–820.
- Pingault, N. et al. (2009). "Produits phytosanitaires et protection intégrée des cultures: l'indicateur de fréquence de traitement". In: *Notes et études socio-économiques* 32, pp. 61–94.
- Pitt, M. M. (1983). "Farm-level fertilizer demand in Java: a meta-production function approach". In: *American Journal of Agricultural Economics* 65.3, pp. 502–508.
- Poulette, B. (2018). "Enquête sur la structure des exploitations agricoles. En 2016, des exploitations moins nombreuses mais plus grandes". In:
- Rabiner, L. R. (1989). "A tutorial on hidden Markov models and selected applications in speech recognition". In: *Proceedings of the IEEE* 77.2, pp. 257–286.
- Rader, T. (1968). "Normally, Factor Inputs Are Never Gross Substitutes". In: *Journal of Political Economy* 76.1, pp. 38–43.
- Reidsma, P. et al. (2010). "Adaptation to climate change and climate variability in European agriculture: The importance of farm level responses". In: *European Journal of Agronomy* 32.1. Cropping Systems Design: new methods for new challenges, pp. 91–102.
- Renaud-Gentié, C., S. Burgos, and M. Benoît (2014). "Choosing the most representative technical management routes within diverse management practices: Application to vineyards in the Loire Valley for environmental and quality assessment". In: *European Journal of Agronomy*.

- Renner, S., J. Sauer, and N. El Benni (2021). "Why considering technological heterogeneity is important for evaluating farm performance?" In: *European Review of Agricultural Economics* 48.2, pp. 415–445.
- Ridier, A. et al. (2013). "The role of risk aversion and labor constraints in the adoption of low input practices supported by the CAP green payments in cash crop farms". In: *Review of Agricultural and Environmental Studies-Revue d'Etudes en Agriculture et Environnement (RAEStud)* 94.906-2016-71162, pp. 195–219.
- Rogers, E. M. (1962). *Diffusion of innovations*. New York: Free Press.
- Rogers, E. M. (2010). *Diffusion of innovations*. Simon and Schuster.
- Rolland, B. et al. (2003). "Des itinéraires techniques à bas niveaux d'intrants pour des variétés rustiques de blé tendre: une alternative pour concilier économie et environnement". In: *Le courrier de l'environnement de l'INRA* 49.49, pp. 47–62.
- Roth, G. et al. (1984). "Effect of Management Practices on Grain Yield, Test Weight, and Lodging of Soft Red Winter Wheat 1". In: *Agronomy journal* 76.3, pp. 379–383.
- Ruud, P. A. (1991). "Extensions of estimation methods using the EM algorithm". In: *Journal of Econometrics* 49.3, pp. 305–341.
- Ryan, B. and N. C. Gross (1943). "The diffusion of hybrid seed corn in two Iowa communities". In: *Rural sociology* 8.1, p. 15.
- Saha, A., C. R. Shumway, and A. Havenner (1997). "The economics and econometrics of damage control". In: *American Journal of Agricultural Economics* 79.3, pp. 773–785.
- Savary, S. et al. (2019). "The global burden of pathogens and pests on major food crops". In: *Nature ecology & evolution* 3.3, pp. 430–439.
- Schepers, J. (2016). "On regression modelling with dummy variables versus separate regressions per group: Comment on Holgersson et al." In: *Journal of Applied Statistics* 43.4, pp. 674–681.
- Schwiebert, J. (2015). "Estimation and interpretation of a Heckman selection model with endogenous covariates". In: *Empirical Economics* 49.2, pp. 675–703.
- Scow, K. et al. (1994). "Transition from conventional to low-input agriculture changes soil fertility and biology". In: *California Agriculture* 48.5, pp. 20–26.
- Serra, T., D. Zilberman, and J. M. Gil (2008). "Differential uncertainties and risk attitudes between conventional and organic producers: the case of Spanish arable crop farmers". In: *Agricultural Economics* 39.2, pp. 219–229.
- Seufert, V. and N. Ramankutty (2017). "Many shades of gray—The context-dependent performance of organic agriculture". In: *Science advances* 3.3, e1602638.
- Seufert, V., N. Ramankutty, and J. A. Foley (2012). "Comparing the yields of organic and conventional agriculture". In: *Nature* 485.7397, pp. 229–232.
- Sexsmith, J. and U. Pittman (1963). "Effect of Nitrogen Fertilizers on Germination and Stand of Wild Oats". In: *Weeds* 11.2, pp. 99–101.
- Shiferaw, B. et al. (2014). "Adoption of improved wheat varieties and impacts on household food security in Ethiopia". In: *Food policy* 44, pp. 272–284.

- Skevas, T., A. O. Lansink, and S. Stefanou (2013). "Designing the emerging EU pesticide policy: A literature review". In: *NJAS-Wageningen Journal of Life Sciences* 64, pp. 95–103.
- Suri, T. (2006). *Selection and Comparative Advantage in Technology Adoption*. Tech. rep. Yale University, Economic Growth Center.
- (2011). "Selection and comparative advantage in technology adoption". In: *Econometrica* 79.1, pp. 159–209.
- Tarde, G. de (1903). *The laws of imitation*. H. Holt and Company.
- Teklewold, H. et al. (2013). "Cropping system diversification, conservation tillage and modern seed adoption in Ethiopia: Impacts on household income, agrochemical use and demand for labor". In: *Ecological Economics* 93, pp. 85–93.
- Tilman, D. et al. (2002). "Agricultural sustainability and intensive production practices". In: *Nature* 418.6898, p. 671.
- Useche, P., B. L. Barham, and J. D. Foltz (2009). "Integrating Technology Traits and Producer Heterogeneity: A Mixed-Multinomial Model of Genetically Modified Corn Adoption". In: *American Journal of Agricultural Economics* 91.2, pp. 444–461.
- Uthes, S. et al. (2019). "Costs, quantity and toxicity: Comparison of pesticide indicators collected from FADN farms in four EU-countries". In: *Ecological Indicators* 104, pp. 695–703.
- Vanloqueren, G. and P. Baret (2009). "How agricultural research systems shape a technological regime but locks out agroecological innovations". In: *Elsevier. Research Policy* 38, pp. 971–983.
- Wei, G. C. and M. A. Tanner (1990). "A Monte Carlo implementation of the EM algorithm and the poor man's data augmentation algorithms". In: *Journal of the American statistical Association* 85.411, pp. 699–704.
- White, H. (2014). *Asymptotic theory for econometricians*. Academic press.
- Wilson, C. (1999). "Cost and policy implications of agricultural pollution, with special reference to pesticides". PhD thesis. University of St Andrews.
- Wilson, C. and C. Tisdell (2001). "Why farmers continue to use pesticides despite environmental, health and sustainability costs". In: *Ecological economics* 39.3, pp. 449–462.
- Wolfson, R. J. (1958). "An econometric investigation of regional differentials in American agricultural wages". In: *Econometrica: Journal of the Econometric Society*, pp. 225–257.
- Wooldridge, J. M. (2002). "Inverse probability weighted M-estimators for sample selection, attrition, and stratification". In: *Portuguese Economic Journal* 1.2, pp. 117–139.
- (2009). "On estimating firm-level production functions using proxy variables to control for unobservables". In: *Economics letters* 104.3, pp. 112–114.
- (2010). *Econometric analysis of cross section and panel data*. MIT press.
- (2015). "Control function methods in applied econometrics". In: *Journal of Human Resources* 50.2, pp. 420–445.

- Wu, J. and B. A. Babcock (1998). "The choice of tillage, rotation, and soil testing practices: Economic and environmental implications". In: *American Journal of Agricultural Economics* 80.3, pp. 494–511.
- Zellner, A., J. Kmenta, and J. Dreze (1966). "Specification and estimation of Cobb-Douglas production function models". In: *Econometrica: Journal of the Econometric Society*, pp. 784–795.
- Zhengfei, G. et al. (2005). "Damage control inputs: a comparison of conventional and organic farming systems". In: *European Review of Agricultural Economics* 32.2, pp. 167–189.
- Zhengfei, G. et al. (2006). "Integrating agronomic principles into production function specification: a dichotomy of growth inputs and facilitating inputs". In: *American Journal of Agricultural Economics* 88.1, pp. 203–214.

Titre : Modélisation micro-économétrique des choix de pratiques de production et des utilisations d'intrants chimiques des agriculteurs : une approche par les fonctions de production latentes.

Mots clés : Fonction de production (duale et primale), pesticides, biais de sélection, modèle à changement de régime endogène, modèle de Markov caché.

Résumé : La notion d'itinéraire technique est une notion agronomique qui nous permet d'appréhender l'imbrication entre les rendements objectifs et les niveaux d'utilisation d'intrants associés. Dès lors, on peut admettre qu'à différents types d'itinéraires techniques correspondent différentes fonctions de production. Modéliser ces différentes fonctions est une des clés pour mieux comprendre la dépendance de certaines pratiques culturales aux pesticides et de ce fait constitue un enjeu majeur pour concevoir les futures politiques publiques.

Intégrer cette notion d'itinéraire technique nécessite de tenir compte de l'interdépendance entre le choix de ces pratiques, leur rendement et les utilisations d'intrants associées. Pour ce faire, on considère des modèles à changement de régime endogène qui permettent de contrô-

ler des biais de sélection. Lorsque ces pratiques sont inobservées, on définit la séquence de choix comme un processus Markovien. Le modèle résultant nous permet de recouvrir les pratiques culturales, leurs niveaux de rendement et d'utilisation d'intrants ainsi que la dynamique de choix des dites pratiques. Lorsque ces pratiques sont observées, on décide de considérer un modèle primal afin de pouvoir vérifier le rôle différencié des pesticides et évaluer l'effet des politiques publiques conjointement sur les rendements et les niveaux d'utilisation d'intrants chimiques.

En bref, cette thèse vise à donner des outils pour évaluer au mieux les effets des politiques agro-environnementales sur les utilisations de pesticides, les rendements et les choix de pratiques culturales des agriculteurs.

Title: Modelling farmers' production choices and chemical inputs demand with a latent function approach

Keywords: Production function (dual and primal), pesticides, selection bias, endogenous regime switching model, hidden Markov model.

Abstract: Cropping management practices is an agronomic notion grasping the interdependence between targeted yield and input use levels. Subsequently, one can legitimately assume that different cropping management practices are associated to different production functions. To better understand pesticide dependence – a key point to encourage more sustainable practices – one have to consider modelling cropping management practices specific production functions.

Because of the inherent interdependence between those practices and their associated yield and input use levels, we need to consider endogenous regime switching models.

When unobserved, the sequence of cropping management practices choices is considered as a Markovian process. From this modelling framework we can derive the cropping management choices, their dynamics, their associated yield and input use levels. When observed, we consider primal production functions to see how yield responds differently to input uses based on the different cropping management practices. Thus, we can assess jointly the effect of a public policy on input use and yield levels.

In a nutshell, in this PhD we are aiming at giving some tools to evaluate the differentiated effect of agri-environmental public policies on production choices and on the associated yield and input use levels.