16èmes Journées de Recherchces en Sciences Sociales, Clermont-Ferrand, 15 et 16 décembre 2022

Intégrer les itinéraires techniques latents aux fonctions de production agricoles : un modèle de Markov caché à paramètres aléatoires

Esther Devilliers(1), Obafemi Philippe Koutchade(2), Alain Carpentier(2)

(1) UMR BETA, INRAE, 54000, Nancy, France.

(2) UMR SMART, INRAE, 35000, Rennes, France.

Auteur de correspondance : esther.devilliers@inrae.fr

Résumé. Pour estimer la réponse des utilisations de pesticides à des changements de prix dans un contexte de technologie hétérogène, cet article intègre la notion agronomique d'itinéraire technique dans les fonctions de production. Généralement, les itinéraires techniques sont des caractéristiques latentes dans les données à disposition des économistes agricoles. La dynamique de choix d'itinéraire technique d'un agriculteur est modélisée à l'aide d'une chaîne de Markov cachée. Par ailleurs, pour contrôler de l'hétérogénéité inobservée entre agriculteurs, nous utilisons une fonction de production à paramètres aléatoires. Nous utilisons l'algorithme SAEM-MCMC pour estimer notre modèle de Markov caché mixte à probabilités de transition hétérogènes. Cette approche nous permet de distinguer trois types d'itinéraires techniques parmi les cultivateurs de blé français situés dans le département de la Marne : des itinéraires très haut-rendement, intermédiaires et bas-intrants.

Mots-clés : Techologie latente, modèle de Markov caché, micro-économétrie de la production agricole.

Accounting for Latent Cropping Management Practices Choices in Crop Production Models: a random parameter Hidden Markov Model Approach

Abstract. In this article, we account for cropping management practices (CMPs) in economists' production functions to evaluate pesticide uses responsiveness to price changes in a context of heterogenous technology. CMPs being latent in most economists' data sets, we consider a hidden Markov model to describe the dynamics of farmer's CMP choice. We also account for farmers' unobserved heterogeneity by considering a random parameter model for our production function. We use the SAEM-MCMC algorithm to estimate our mixed hidden Markov model with heterogeneous transition probabilities. An illustration on French winter wheat producers of La Marne area allows us to distinguish between high-yielding, intermediate and low-input practices.

Keywords: Latent technology, mixed hidden-Markov model, micro-econometrics of agricultural production.

Classification JEL : C13, C24, Q12.

1 Introduction

Regulation of pollutions due to the use of chemical inputs, pesticides in particular, in agricultural production is a major policy objective in the European Union (EU). 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. But, econometric results tend to demonstrate that farmers' pesticide uses generally display very limited responsiveness to pesticide prices, implying that incentive taxes on pesticides would have limited effects on their uses together with significant effects on farmers' incomes (see, *e.g.*, Aubertot et al., 2007; Böcker and Finger, 2017; Frisvold, 2019; Skevas et al., 2013).

Recent pesticide demand own price elasticity estimates, however, are based on cross-section or on short panel data and generally rely on standard, and often dual, production choice models (Böcker and Finger, 2017). These incentives these estimates generally rely on limited input price variations. Moreover, estimates ignore that farmers can switch from pesticide dependent to pesticide saving crop production technologies, thereby widening the scope of pesticide use responses to price. Indeed, according to agricultural scientists, farmers cannot significantly reduce their chemical input uses without severe economic losses unless they change their cropping management practices (CMPs). This suggests that considering CMPs – their choices by farmers and their effects on yield and input use levels – in crop production choices is likely to be useful for revisiting pesticide uses of farmers.

Agronomists' CMP concept is closely related to economists' (crop) production technology concept. Agronomists define a CMP as an ordered sequence of operations – *e.g.*, – aimed to produce a given crop. Considered operations include tillage operations, sowing date and density of a given variety, pesticide and fertilizer application dates and quantities. CMPs for a given crop differ according to their target yield levels and/or to their reliance on specific categories of inputs. Of course, targeted yield levels and input uses are closely related, these features of CMPs being consistently determined based on the considered crop features. For instance, conventional CMPs for wheat in France are high yielding. They rely on productive seed varieties, on high levels of chemical input uses – *i.e.* fertilizers and pesticides – as well as on specific techniques, including early and dense sowing.[1] Organic CMPs prohibit uses of mineral fertilizers and synthetic pesticides. They make use of hardy seeds varieties (which are often more resistant to pest and disease but less productive than varieties selected for their productivity), organic fertilizers or non-synthetic pesticides. They, however, achieve lower yield levels than conventional CMPs. Yield levels of organic wheat production amount to around half of those of conventional production. In the 1990s, French agronomists conceived and experimented low-input CMPs, which is of special interest in this study, for wheat production as a middle ground between conventional and organic CMPs (Loyce et al., 2008, 2012; Meynard, 1985; Meynard, 1991). These CMPs basically trade-off small reductions in yield levels for significant reductions in pesticide uses, in fungicide uses in particular. They rely on hardy seed varieties and avoid techniques that tend to enhance pest and weed pressures, such as early and dense sowing.

Unfortunately, targeted yield levels and cropping techniques characterizing CMPs are unobserved in datasets used by agricultural production economists for investigating farmers' production choices. Also, farmers' yield and chemical input use levels are impacted by factors and events that make their direct comparison across farms and years largely irrelevant for identifying CMPs.[2] Similarly, climatic events or pest and disease infestations impact yields and input uses in ways that significantly differ across farms and years. This across farm and year heterogeneity has significant effects on crop production and on farmers' choices even within areas of limited size (*e.g.* Koutchadé et al., 2018, 2021).

This article proposes a micro-econometric model featuring CMP choices and their effects on crop chemical input uses and yields. The proposed model considers the yield and chemical input use levels of a given crop and assumes that the observed yields and input uses stem from a set CMPs that are not observed. The considered model is to be estimated based on a panel dataset containing

[1]Early sowing increases the photosynthesis duration but also increases crop exposure to pest and diseases. Dense sowing increases yield levels by augmenting plant number by unit of land but also increases fertilization needs as well as crop protection needs. High nitrogen fertilization rates tend to increase crop susceptibility to diseases and favor weeds when these are not suitably controlled by tillage operations or herbicide uses. Dense vegetal covers favor pest and disease spreading.

[2]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 type. High-yielding CMPs may target 9.5 t/ha of wheat in a good plot while they may only target 7.5 t/ha in a plot with poor soil quality.

cost accounting data of a large sample of farmers. The main objectives of this empirical application are to uncover and to characterize the main CMPs used by the sampled farmers on the one hand, and to assess the effects of major economic drivers on the choices of the identified CMPs on the other hand.

The micro-econometric model we consider for modelling yield and input use levels is a hidden Markov model featuring farm specific random parameters and a Markov transition probability matrix explicitly incorporating the effects of the considered CMP economic returns. Several features of the considered process and data lie at the root of our choosing this modelling framework. (a) Following Fémenia and Letort (2016), we take advantage of the conceptual similarity between agronomists' CMPs and economists' production technologies. Because CMP choices are unobserved, we consider that observed yield and input use levels are generated by a set CMP specific yield supply and input demand models that are latent in our data. (b) We assume that farmers choose the CMP they use at the beginning of the cropping season, thereby assuming that CMP choices are short run ones. Indeed, the choice of a CMP involves neither large investment costs nor significant inter-temporal trade-offs. It shares much more similarities with standard crop variety choices (*e.g.*, Michler et al., 2018; Suri, 2011), than with irrigation technology choices (*e.g.*, Genius et al., 2014). (c) Nevertheless, agronomists report that farmers tend to keep on using the same CMP unless sufficient changes in the regulatory or economic environments lead them to reconsider their CMP. Indeed, CMP changes may entail intangible costs, such as learning costs when farmers consider CMPs they are not familiar with. Accordingly, we assume that farmers' current CMP choices can depend on their previous CMP choices and, thereby, model farmers' CMP choice as a multivariate discrete Markovian process. This implies that our model of yield and input use levels is a so-called hidden Markov model (HMM). (d) In order to investigate the effects of the main economic drivers of farmers' CMP choices, we derive the functional form of the Markovian transition probability of our HMM by considering CMP choices as expected return maximizing discrete choices involving switching costs. (e) Following Koutchadé et al. (2018, 2021), most of our model parameters are considered as farm specific random parameters to account for unobserved heterogeneity in farm production condition (*e.g.*, machinery endowment, soil quality) as well as in farmer skills or motivations (*e.g.*, pro-environment attitude).

Our empirical application is based on cost accounting panel data covering a sample of more than 1300 French wheat producers from 1998 to 2014. Sampled farmers are located in *La Marne*, which is a small French territorial division (called a *département*) located in eastern France (in the *Champagne région*). Besides its being highly productive for arable crops, La Marne is of special interest regarding CMPs and pesticide uses. The low input CMPs described above were experimented on-farm in this area during the mid-90s (Loyce and Meynard, 1997). Our results confirm the relevance of considering CMPs when analyzing the considered farmer choices. In particular, our HMM uncovers three CMP types, including a pesticide saving CMP type featuring characteristics close to the ones of the low-input CMPs developed by French agronomists. Yet, our results also suggest that CMPs currently available for wheat production in France do not enable farmers to significantly reduce their uses of pesticides without incurring significant economic losses. The pesticide saving CMPs uncovered in our application either entail reductions in pesticide uses that are too limited or imply losses in yields that are too large for these CMPs to be widely adopted. For instance, according to our results the low-input CMPs were adopted by up to 15% of the sampled farmers before 2007. This adoption rate felt down to 5% after 2007, that is to say after wheat prices raised by around 65% while pesticide prices remained basically constant.

Goldfeld and Quandt (1973) and Hamilton (1989) introduced HMM in econometrics, in macroeconometrics in particular. Most econometric studies considering this HMM framework consider time series, especially financial data (*e.g.*, Bonomo and Garcia, 1996; Chauvet and Hamilton, 2006; Rydén et al., 1998).[3] CMP choice sequences varying across farms in our panel data, our modelling framework is related to HMMs developed in the statistics literature, econometrics excluded to our knowledge, for longitudinal data (*e.g.*, Altman, 2007; Maruotti, 2011). Our model featuring random parameters in both its transition probability matrix and its yield and input use models, it can also be considered as a so-called mixed HMM with heterogeneous probabilities (*e.g.*, Lavielle, 2018). Also, the two main components of our model – *i.e.*, its dynamic CMP choice sub-model and its yield and input use levels

[3] Yet, their use in econometrics is still marginal compared to other fields as genomics. New technologies in DNA sequencing allow scientists to collect large DNA data. Hidden Markov models are then used to predict protein structure and function (Eddy, 1996; Henderson et al., 1997; Prestat et al., 2014) or to model the DNA copy number change across the genome (Manogaran et al., 2018) for instance. More generally, HMM is a very useful tool in pattern recognition among big data sets (see, *e.g.*, Elmezain et al., 2009; Nguyen et al., 2005; Varga and Moore, 1990). Yet, the aforementioned applications mostly involve individual stochastic processes.

CMP specific sub-models – share common random parameters (and, more, generally contain random parameters that can be correlated). This implies that our model exhibits the distinctive properties of the endogenous Markov switching models considered by Kim et al. (2008) or Hwu et al. (2019).

Estimating the random parameter HMM considered in this article is challenging. As suggested by Lavielle (2018), we solve this Maximum Likelihood estimation problem by nesting a Baum-Welch (forward-backward) algorithm in a specifically designed Stochastic Approximate Expectation-Maximization (SAEM) algorithm. The Baum-Welch (or forward-backward) algorithm is a convenient way to compute the CMP choice probabilities that are part of the likelihood function of our model (*e.g.*, Maruotti, 2011; Welch, 2003). Econometricians generally rely on Simulated Maximum Likelihood estimators for parametric models involving random parameters. But other statisticians often prefer Expectation-Maximization (EM) type algorithms, which were originally developed by Dempster et al. (1977) for similar problems. Indeed, EM type algorithms are particularly well suited for maximizing the log-likelihood functions of models involving missing variables, such as latent CMP choices and random parameters in our model. EM type algorithms basically solve complicated ML problems by iteratively solving sequences of much simpler problems. Each iteration of an EM type algorithm involves an expectation (E) step that consists of integrating a conditional expectation, and a maximization (M) step that updates the estimate of the parameter of interest. The Monte Carlo EM (MCEM) algorithms proposed by Wei and Tanner (1990) extend the original deterministic EM algorithms for handling cases in which the E step integration problem needs to be solved with simulation methods. The SAEM algorithms proposed by Delyon et al. (1999), which rely on stochastic approximations for solving the E step, are computationally efficient alternative to MCEM algorithms, especially when the probability distributions involved in the likelihood function of the model belong to the exponential family as in our case (*e.g.*, Kuhn and Lavielle, 2005; Lavielle and Mbogning, 2014).

The random parameter HMM model proposed in this article for analyzing agricultural production choices enables us to obtain original results of production practices in France based on a rich panel dataset of cost accounting observations. In particular, our estimation results reveals that farmers in La Marne mostly rely on relatively high yielding CMPs, which are intensive in chemical inputs, for producing wheat. Yet, our results also reveals that a small fraction of farmers use pesticide saving

CMPs. We expected to uncover such CMPs because very close low input CMPs were experimented by agronomists in the considered area a few years before the period covered by our data. Importantly, our estimates tend to demonstrates that the CMP choices of the considered farmers display significant persistence over time but also respond to economic incentives, albeit to a limited extent. These results shed a new light on the economic and technological dependence of arable crop production in the European Union.

Several other models features components and properties of the random parameter HMM model considered in this article. But, to our knowledge, this model is the first to combine all these components and properties. (a) Considering latent technologies is now common practice in the stochastic production frontier literature. Latent class stochastic production frontier models were proposed by Orea and Kumbhakar (2004) and Greene (2005), and then used by Alvarez and Corral (2010), Martinez Cillero et al. (2018), Renner et al. (2021) or Dakpo et al. (2021). (b) Most of the parameters of our model are farm specific random parameters, as in the micro-econometric multi-crop models of Koutchadé et al. (2018, 2020) or in the input demand model of Los et al. (2021). (c) We assume that farmers' CMP choices can be modelled as Markovian processes. This approach provides a solution to an issue encountered when considering standard latent class models with panel data. Standard latent class models can only describe two technology choice patterns (*e.g.*, Renner et al., 2021); producers either stick to the same technology along the observation period or "freely" choose the technology they use each year. Our dynamic CMP choice model can describe CMP choice patterns lying between these two polar cases; farmers can switch from a CMP to another, but their CMP choices are (more or less softly) constrained by switching costs. (d) Following the seminal work of Griliches (1957), we explicitly define the expected economic returns of the CMPs under consideration as key drivers of farmers' CMP choices. This casts our study into the considerable economic literature dealing with the adoption and the diffusion of agricultural production technologies. Yet, the production technologies are observed in the data used by most the micro-econometric analyses considered in this literature.[4]

[4] These analyses 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 drivers such as learning processes and uncertainties (*e.g.*, Chavas and Nauges, 2020; Foster and Rosenzweig, 2010; Marra and Pannell, 2003), heterogeneity in returns to adoption (*e.g.*, Michler et al., 2018; Suri, 2011) or labor constraints (*e.g.*, Fernandez-Cornejo et al., 2005). A few studies 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).

The rest of the article is organized as follows. The second section of the article discusses in more details low-input 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 a 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 from the random parameter hidden Markov model. Lastly, we discuss the obtained results and provide some concluding remarks.

2 Agronomic principles and brief history of "Low Input" CMPs

As stated before, we are interested in specific CMPs: the low-input CMPs (LI-CMPs) and the high-yielding CMPs (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 et al., 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 et al., 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, as developed by INRA starting in the mid 1980s, 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. 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.

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). 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). Yet, the 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. If yield levels obtained with LI-CMPs are on average 10% lower than those obtained with HY-CMPs, farmers benefit from multiple input savings and thus cost reduction. Indeed, 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%. Plus, due to the lower sowing densities in LI-CMPs seed uses decrease by around 50% when using these 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, we expect that, given the observed economic conditions (low gain prices while stable prices for fungicides and insecticides), the adoption of LI-CMPs was favored from the late 1990s to 2006. 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 A Random Parameter Hidden Markov Model for modelling production choices accounting for CMPs

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 et al., 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 evolue 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.

3.2 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 $J = \{1, ..., J\}$ is the considered set of inputs. As our dataset is a unbalanced panel, we also need to define $H_i = \{t(i), ..., T(i)\}$ as the observation period of farmer i, where $i = 1, ..., N$ and $H_i \subseteq \{1, ..., T\}$. We assume that farmers can produce wheat by using a CMP among the C ones collected in set $C = \{1, ..., C\}$. 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^c_{it} denotes the wheat level obtained by farmer i considering that this farmer used CMP c in year t. Vector $\mathbf{x}^c_{it} = (x^c_{j,it} : j \in J)$ denotes the corresponding input use levels. The model chosen for the $(y^c_{it}, \mathbf{x}^c_{it})$ vectors is given by

$$\begin{cases} y^c_{it} = b^c_{y,i} + d_{y,t,0} + \boldsymbol{\delta}'_{y,0}\mathbf{z}_{it} + \varepsilon^c_{y,it} \\ \mathbf{x}^c_{it} = \mathbf{b}^c_{x,i} + \mathbf{d}_{x,t,0} + \boldsymbol{\Delta}_{x,0}\mathbf{z}_{it} + \boldsymbol{\varepsilon}^c_{x,it} \end{cases} \quad \text{, for } c \in C, \tag{1}$$

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.[5] 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 $\boldsymbol{\Delta}_0 = (\boldsymbol{\delta}'_{y,0}, \boldsymbol{\Delta}_{x,0})$, 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.[6] Farm and CMP specific terms $\mathbf{b}^c_i = (b^c_{y,i}, \mathbf{b}^c_{x,i})$, where $\mathbf{b}^c_{x,i} = (b^c_{x,j,i} : j \in J)$, account for the effects of CMP and production conditions on wheat yields and input uses. Term $b^c_{y,i}$ is designated, for short, as the wheat target yield level of CMP c as this practice is implemented by farmer i. Similarly, term $b^c_{x,j,i}$ is designated as the requirement in input j of CMP c. Error terms $\boldsymbol{\varepsilon}^c_{it} = (\varepsilon^c_{y,it}, \boldsymbol{\varepsilon}^c_{x,it})$ 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 $\mathbf{b}^c_i$, $\mathbf{z}_{it}$ and $\boldsymbol{\varepsilon}^c_{it}$ are mutually independent.[7] We also assume that error terms $\boldsymbol{\varepsilon}^c_{it}$ are independent across farms and years. These assumptions imply that vectors $\mathbf{q}_{it}$ are independent across time conditionally on $\mathbf{b}^c_i$ and $\mathbf{z}_{it}$. Finally we assume that $\boldsymbol{\varepsilon}^c_{it}$ is normally distributed, with $\boldsymbol{\varepsilon}^c_{it} \sim N(\mathbf{0}, \boldsymbol{\Sigma}^c_0)$.

[5] For 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.

[6] These homogeneity assumptions are admittedly restrictive as latent vector netput levels $\mathbf{q}^c_{it} = (y^c_{it}, \mathbf{x}^c_{it})$ 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 simplifies our CMP choice model. Third, and more importantly, the effects of CMPs are captured in other parts of the model of $\mathbf{q}^c_{it}$.

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

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 $b^c_{y,i}$ and $b^c_{x,j,i}$ based on simple recursive schemes, $b^c_{y,i} = a^c_{y,i} b^{c-1}_{y,i}$ and $b^c_{x,j,i} = a^c_{x,j,i} b^{c-1}_{x,j,i}$. It implies that $b^c_{y,i}$ and $b^c_{x,j,i}$ are given by the following simple formulae:

$$b^c_{y,i} = b^1_{y,i} \prod\nolimits_{d=2}^{c} a^d_{y,i} \qquad , \text{ for } c \in C_a,$$

and

$$b^c_{x,j,i} = b^1_{x,j,i} \prod\nolimits_{d=2}^{c} a^d_{x,j,i} \qquad , \text{ for } c \in C_a \text{ and } j \in J,$$

where $C_a = \{2, ..., C\}$. The conditions stating that $b^1_{y,i} \geq 0$, $a^c_{y,i} \in [0,1]$, $b^1_{x,j,i} \geq 0$ and $a^c_{x,j,i} \in [0,1]$, for $c \in C_a$ and $j \in J$, guarantee that expected yield and chemical input use levels $b^c_{y,i}$ and $b^c_{x,j,i}$ 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 $\boldsymbol{\gamma}_i = (\mathbf{b}^1_i, \mathbf{a}_{y,i}, \mathbf{a}_{x,i})$, where $\mathbf{a}_{y,i} = (a^c_{y,i} : c \in C_a)$, $\mathbf{a}^c_{x,i} = (a^c_{x,j,i} : j \in J)$ and $\mathbf{a}_{x,i} = (\mathbf{a}^c_{y,i} : c \in C_a)$.[8] 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:

$$f(\mathbf{q}_{it} | r_{it} = c, \boldsymbol{\gamma}_i, \mathbf{z}_{it}) = \varphi(\mathbf{q}_{it} - \mathbf{b}^c_i - \mathbf{d}_{t,0} - \boldsymbol{\Delta}_0 \mathbf{z}_{it}; \boldsymbol{\Sigma}^c_0), \tag{2}$$

where $\varphi(\mathbf{a}; \boldsymbol{\Xi})$ is the probability distribution function of $N(\mathbf{0}, \boldsymbol{\Xi})$ 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 $\boldsymbol{\gamma}_i$).

[8]Interestingly, 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 $i = 1, ..., N$. 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.

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:

$$\mathbf{q}_{it} = \sum\nolimits_{c \in C} 1(r_{it} = c)\mathbf{q}_{it}^{c} = \sum\nolimits_{c \in C} r_{it}^{c}\mathbf{q}_{it}^{c}, \tag{3}$$

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.[9] 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 $\boldsymbol{\gamma}_i$ 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:

$$\pi_{it}^c = E[w_{y,t}|\omega_{it}]b_{y,i}^c - \mathbf{w}'_{x,t}\mathbf{b}_{x,i}^c + \kappa_{it}, \tag{4}$$

where

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

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 inves-

[9] In 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.

tigating 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 et al., 2018, 2020). It implies that $\pi_{it}^{c} = w_{y,t-1}b_{y,i}^{c} - \mathbf{w}'_{x,t}\mathbf{b}_{x,i}^{c} + \eta_{it}$.[10]

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 $\boldsymbol{\pi}_{it}$. We use a first order Markov chain so we have that $P_i[r_{it} = d|r_{it-1}, ..., r_{it(i)}, \boldsymbol{\pi}_{it}] = P_i[r_{it} = d|r_{it-1}, \boldsymbol{\pi}_{it}]$.[11]

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 $\boldsymbol{\pi}_{it}$ and of implicit CMP switch costs. Three main types of switch costs can be defined. First, expected returns $\boldsymbol{\pi}_{it}$ 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 $\boldsymbol{\alpha}_i$ stack all the farm specific random parameters of the entire model and function $p(d|c, \boldsymbol{\alpha}_i, \mathbf{w}_t)$ 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 $\boldsymbol{\alpha}_i$ and on expected price levels $\mathbf{w}_t = (w_{y,t-1}, \mathbf{w}_{x,t})$. This probability function is defined by $p(d|c, \boldsymbol{\alpha}_i, \mathbf{w}_t) = P[r_{it} = d|r_{it-1} = c, \mathbf{w}_t, \boldsymbol{\alpha}_i]$ and its assumed

[10] Considering 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.

[11] The underlying assumption being that the last CMP choice, r_{it-1}, sums up the information content of the CMP choice history, $(r_{it-1}, ..., r_{it(i)})$, that is relevant for modelling CMP choice r_{it} given $\boldsymbol{\pi}_{it}$.

functional form is given by:

$$p(d|c, \boldsymbol{\alpha}_i, \mathbf{w}_t) = \frac{\exp\left(\rho_i(\pi_{it}^d - \mu_i^{d|c})\right)}{\sum_{k \in C} \exp\left(\rho_i(\pi_{it}^k - \mu_i^{k|c})\right)} \qquad , \text{ for } (c,d) \in C \times C. \tag{5}$$

The functional form of transition probability function $p(d|c, \boldsymbol{\alpha}_i, \mathbf{w}_t)$ is that of a (mixed) 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 $t-1$.[12] 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 (5), we need to assume that the elements of error term vector $\mathbf{e}_{it}^{|c} = (e_{it}^{d|c} | d \in C)$, 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 (5), farmers' CMP choice depends on the relative expected returns of the CMPs, $\boldsymbol{\pi}_{it}$, as well as on farmer specific switching costs $\boldsymbol{\mu}_i = (\mu_i^{d|c} : (c,d) \in C^2)$.[13]

Equation (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 $p_0(c|\boldsymbol{\alpha}_i, \mathbf{w}_{t(i)})$ denote such probability given $\boldsymbol{\alpha}_i$ and $\mathbf{w}_{t(i)}$. We assume that the functional form $p_0(c|\boldsymbol{\alpha}_i, \mathbf{w}_{t(i)})$ is given by:

$$p_0(c|\boldsymbol{\alpha}_i, \mathbf{w}_{t(i)}) = \frac{\exp\left(\sigma_i(\pi_{it(i)}^c - \eta_i^c)\right)}{\sum_{k \in C} \exp\left(\sigma_i(\pi_{it(i)}^k - \eta_i^k)\right)} \qquad , \text{ for } c \in C. \tag{6}$$

This functional form is chosen to account for the effects of the CMP relative profitability levels $\boldsymbol{\pi}_{it(i)}$.

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

[13] We 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.

It is inspired by the Multinomial Logit probability function associated to the expected profit maximization problem given by $\max_{k\in C}\{\pi^k_{it(i)} - \eta^k_i + (\sigma_i)^{-1}e^k_{it(i)}\}$. Term σ_i is a positive farm specific parameter scaling the effects of error terms $e^k_{it(i)}$ and terms $\boldsymbol{\eta}_i = (\eta^c_i : c \in C)$ capture the effects of farmer specific costs or motives that tend to direct farmers' choice toward particular CMPs.[14] 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^c_i to account for the fact that, depending on when farmer i arrives in the panel, his probability to adopt a specific CMP might changes.[15]

Combining Equations (5) and (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 $P(\mathbf{r}_{(i)}|\boldsymbol{\alpha}_i, \boldsymbol{w}_{(i)})$ 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:

$$\begin{aligned}\ln P_i(\mathbf{r}_{(i)}|\boldsymbol{\alpha}_i) = &\sum\nolimits_{c\in C} r^c_{it(i)} \ln p_0(c|\boldsymbol{\alpha}_i, \mathbf{w}_{t(i)}) \\ &+ \sum\nolimits_{t=t(i)+1}^{T(i)} \sum\nolimits_{c\in C} \sum\nolimits_{d\in C} r^c_{it-1} r^d_{it} \ln p(d|c, \boldsymbol{\alpha}_i, \mathbf{w}_t).\end{aligned} \tag{7}$$

Equation (7) corresponds to the likelihood of the considered CMP choice model for our panel dataset, with $\boldsymbol{\alpha}_i = (\boldsymbol{\gamma}_i, \rho_i, \boldsymbol{\mu}_i, \sigma_i, \boldsymbol{\eta}_i)$. If we assume that error terms $\mathbf{e}_{it} = (\mathbf{e}^{|c}_{it}|c \in C)$ and $\boldsymbol{\varepsilon}_{it}$ are independent, we can easily derives the unconditional likelihood associated to $\mathbf{q}_{it}$.[16]

[14] Condition $\eta^1_i = 0$ is chosen as the normalization constraint for the elements of $\boldsymbol{\eta}_i$.

[15] This point is discussed further in the Discussion/Conclusion section.

[16] Such 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 $\boldsymbol{\pi}_{it}$ is a function of $\boldsymbol{\gamma}_i$, the random parameters characterizing the CMPs. Also, we allow switching cost random parameters $\boldsymbol{\mu}_i$ to be correlated with $\boldsymbol{\gamma}_i$. 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 $\boldsymbol{\alpha}_i$.

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:

$$\ell_i(\boldsymbol{\theta}|\boldsymbol{\alpha}_i) = \sum\nolimits_{c_{it(i)} \in C} .... \sum\nolimits_{c_{iT(i)} \in C} p_0(c_{it(i)}|\boldsymbol{\alpha}_i, \mathbf{w}_{t(i)}) \prod\nolimits_{t=t(i)+1}^{T(i)} p(c_{it}|c_{it-1}, \boldsymbol{\alpha}_i, \mathbf{w}_t) \prod\nolimits_{t=t(i)}^{T(i)} \varphi(\mathbf{u}_{it}^c; \boldsymbol{\Sigma}^{c_{it}}), \tag{8}$$

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:

$$\ell_i(\boldsymbol{\theta}) = \int \ell_i(\boldsymbol{\theta}|\boldsymbol{\varsigma})\varphi(\boldsymbol{\varsigma} - \boldsymbol{\alpha}; \boldsymbol{\Omega})d\boldsymbol{\varsigma}. \tag{9}$$

The computation of likelihood terms $\ell_i(\boldsymbol{\theta})$ is particularly challenging.[17] Hence the use of extensions of the Expectation-Maximization (EM) algorithm of Dempster et al. (1977).[18]

The Stochastic Approximate EM (SAEM) algorithm proposed by Delyon et al. (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

[17] This problem combines two issues. As it stands in Equation (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 (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 $\boldsymbol{\theta}$.

[18] EM 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.

$\boldsymbol{\zeta}_{(i)} = (\mathbf{q}_{(i)}, \mathbf{w}_{(i)}, \mathbf{z}_{(i)})$, (ii) the latent CMP choices sequence $\mathbf{r}_{(i)}$ and (iii) the random parameters vector $\boldsymbol{\alpha}_i$, for $i = 1, ..., N$. 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:

$$\ln L^C(\boldsymbol{\theta}) = \sum\nolimits_{i=1}^{N} \ln \ell_i^C(\boldsymbol{\theta}|\mathbf{r}_{(i)}, \boldsymbol{\alpha}_i), \tag{10}$$

where:

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

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 $\boldsymbol{\zeta}_{(i)}$ evaluated at $\boldsymbol{\theta}^{(n)}$, the last available estimate of $\boldsymbol{\theta}_0$. This consists of computing the conditional expectations $E^{(n)}[\ln \ell_i^C(\boldsymbol{\theta}|\mathbf{r}_{(i)}, \boldsymbol{\alpha}_i)|\boldsymbol{\zeta}_{(i)}]$, for $i = 1, ..., N$.

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

$$E^{(n)}[\ln \ell_i^C(\boldsymbol{\theta}|\mathbf{r}_{(i)}, \boldsymbol{\alpha}_i)|\boldsymbol{\zeta}_{(i)}] \\ = \\ \left\{ \begin{array}{c} E^{(n)}[R(\boldsymbol{\alpha}_i, \mathbf{w}_{(i)})|\boldsymbol{\zeta}_{(i)}] \\ +\sum_{t=t(i)}^{T(i)} \sum_{c\in C} E^{(n)}[s_t(c,d|\boldsymbol{\alpha}_i, \mathbf{w}_{(i)}) \ln \varphi(\mathbf{q}_{it} - \mathbf{b}_i^c - \mathbf{d}_t - \boldsymbol{\Delta}\mathbf{z}_{it}; \boldsymbol{\Sigma}^c)|\boldsymbol{\zeta}_{(i)}] \\ +E^{(n)}[\ln \varphi(\boldsymbol{\alpha}_i - \boldsymbol{\alpha}; \boldsymbol{\Omega})|\boldsymbol{\zeta}_{(i)}] \end{array} \right\}, \tag{11}$$

where term:

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

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

The expectations conditional on the observed data $\boldsymbol{\zeta}_{(i)}$ involved in Equation (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 $\widehat{\boldsymbol{\alpha}}_i^{(n),j}$ denote the considered random draws of $\boldsymbol{\alpha}_i$ and $\widehat{\omega}_i^{(n),j}$ their related weights, for $j = 1, ..., J^{(n)}$, where $J^{(n)}$ denote the draw number considered for iteration n. The conditional expectation of $\ln L^C(\boldsymbol{\theta})$ is approximated by:

$$E^{(n)}[\ln L^C(\boldsymbol{\theta})|\boldsymbol{\zeta}] \\ \simeq \\ \sum_{i=1}^N \left\{ \begin{array}{c} E^{(n)}[R(\boldsymbol{\alpha}_i, \mathbf{w}_{(i)})|\boldsymbol{\zeta}_{(i)}] \\ +\sum_{t=t(i)}^{T(i)} \sum_{c\in C} \sum_{j=1}^{J^{(n)}} \widehat{\omega}_i^{(n),j} s_t(c,d|\widehat{\boldsymbol{\alpha}}_i^{(n),j}, \mathbf{w}_{(i)}) \ln \varphi(\mathbf{q}_{it} - \widehat{\mathbf{b}}_i^{c,(n),j} - \mathbf{d}_t - \boldsymbol{\Delta}\mathbf{z}_{it}; \boldsymbol{\Sigma}^c) \\ +\sum_{j=1}^{J(n)} \widehat{\omega}_i^{(n),j} \ln \varphi(\widehat{\boldsymbol{\alpha}}_i^{(n),j} - \boldsymbol{\alpha}; \boldsymbol{\Omega}) \end{array} \right\}. \tag{12}$$

In our empirical application, we used the SAEM version of this E step and an importance sampling approach for integrating terms $E^{(n)}[\ln \ell_i^C(\boldsymbol{\theta}|\mathbf{r}_{(i)}, \boldsymbol{\alpha}_i)|\boldsymbol{\zeta}_{(i)}]$. 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 $\boldsymbol{\alpha}_i$.[19]

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

$$\max_{(\mathbf{d},\boldsymbol{\Delta},\boldsymbol{\Sigma})} \sum_{i=1}^N \sum_{t=t(i)}^{T(i)} \sum_{c\in C} \sum_{j=1}^{J^{(n)}} \widehat{\omega}_i^{(n),j} s_t(c,d|\widehat{\boldsymbol{\alpha}}_i^{(n),j}, \mathbf{w}_{(i)}) \ln \varphi(\mathbf{q}_{it} - \widehat{\mathbf{b}}_i^{c,(n),j} - \mathbf{d}_t - \boldsymbol{\Delta}\mathbf{z}_{it}; \boldsymbol{\Sigma}^c), \tag{13}$$

[19] The whole SAEM procedure and its explicit forms can be found in Appendix 7.1.

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

On the other hand, solving problem:

$$\max_{(\boldsymbol{\alpha},\mathbf{\Omega})} \sum_{i=1}^{N} \sum_{j=1}^{J(n)} \widehat{\omega}_i^{(n),j} \ln \varphi(\widehat{\boldsymbol{\alpha}}_i^{(n),j} - \boldsymbol{\alpha}; \mathbf{\Omega}), \tag{14}$$

yields $(\boldsymbol{\alpha}^{(n+1)}, \mathbf{\Omega}^{(n+1)})$, the second part of $\boldsymbol{\theta}^{(n+1)}$. 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 7.1 for more details on the estimation procedure).

5 Data

5.1 Data description

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 1.

Table 1: Annual mean (and standard deviation) of arable land area and wheat acreage, from 1999 to 2014

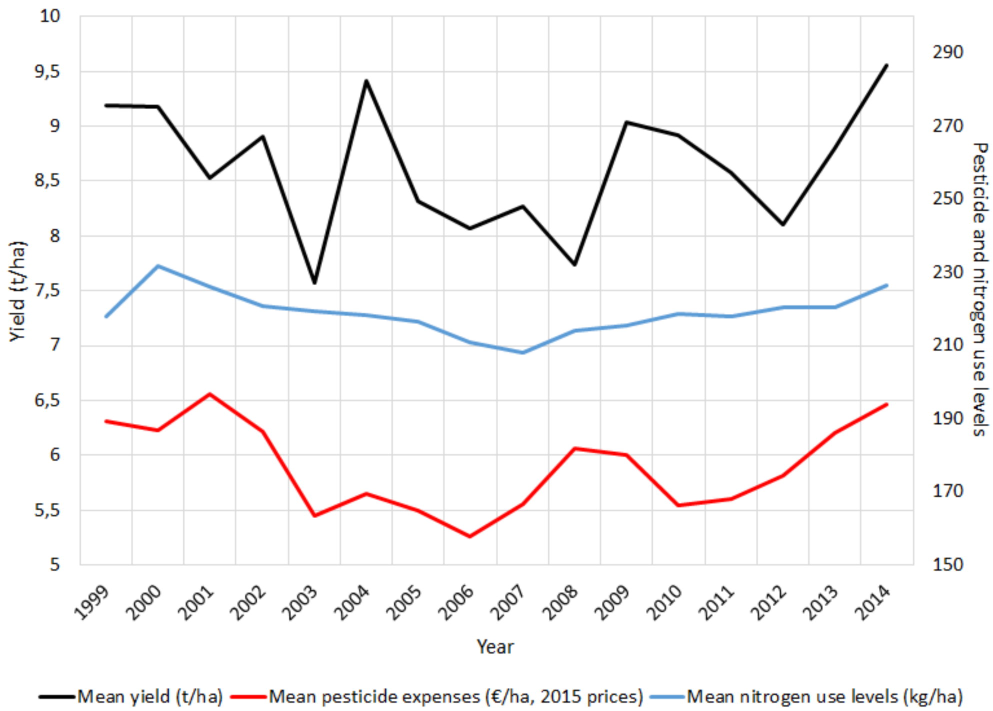
Year	N	Arable land area (in ha)	Wheat acreage (in ha)
1999	678	172.61 *(86.54)*	52.40 *(29.31)*
2000	824	179.11 *(89.92)*	57.42 *(33.41)*
2001	829	176.60 *(90.98)*	53.91 *(30.33)*
2002	959	181.89 *(92.45)*	56.42 *(33.24)*
2003	946	184.96 *(92.36)*	53.74 *(32.49)*
2004	924	185.40 *(93.46)*	55.10 *(33.32)*
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.

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 $\mathbf{d}_{it}$.

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). Mean yield, fertilizer and pesticide expenses are reported in Figure 1.

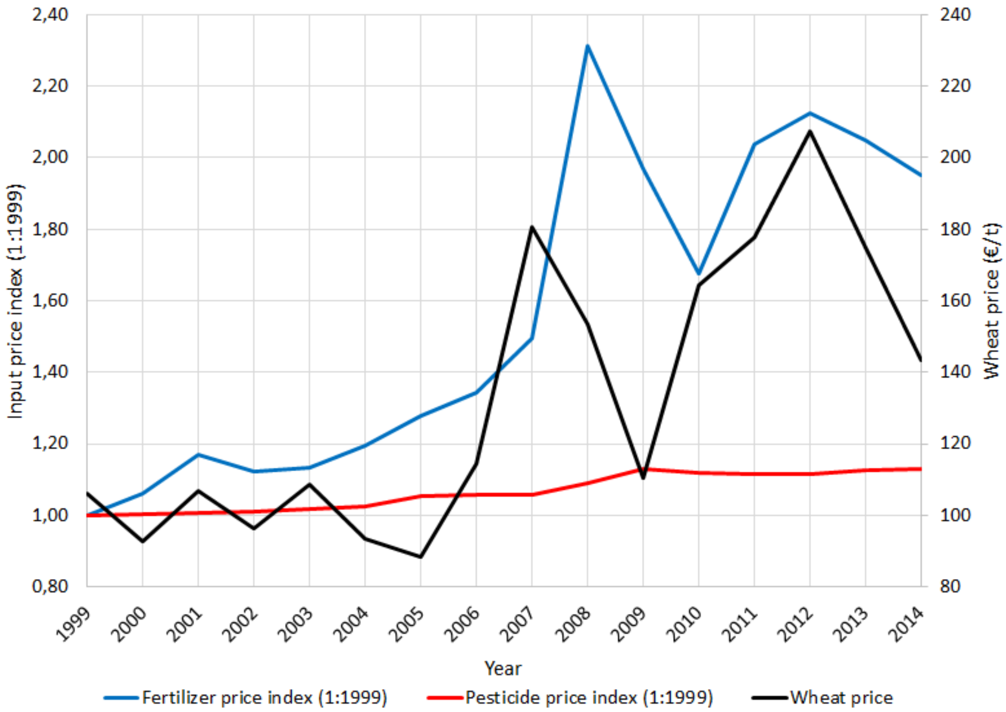
Figure 1: Annual mean yields and input expenses from 1999-2014



Source: CDER data.

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 2 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 2: 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.

5.2 Exploratory analyses

Before imposing our structural framework, we wanted to confirm the coexistence of CMPs in the dataset by considering non-structural approaches. In particular, we consider clustering analysis combined to a latent class model to try and find groups among our dataset based on the observed yield and input uses. Yet, standard cluster analysis and latent class models consider static clusters which is problematic when considering panel data as we do. In that case, either we can perform a cluster analysis for each year separately and suppose that the group belonging is independent between years. Or, we can perform a cluster analysis on all years while assuming that an individual belongs to the same group for the length of the time period considered. Both approaches are unsatisfactory when considering technology adoption. situations generate pitfalls in the case of technology adoption. We choose an in-between solution which is to divide the data set into 4 years sub-panels.[20] As input uses

[20]Sub-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 ourself to this area because effects of exogeneous factors such soil fertility and climate on yields and chemical inputs can be

- 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 et al., 2010). After the principal component analysis, we performed an ascendant hierarchical clustering analysis, and considered the obtained classes as initial value for our latent class model. Detailed results from the latent class model with two and three classes can be found in Appendix 7.2. In what follows, we focus on the main insights we get from this non-structural approach.

First, when considering two classes, we find stable groups across the different subpanels. 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 the "low-input" CMP yield level is lower by 7% in comparison to the "high-yielding" one. As for input expenses, differences are of 5% for nitrogen and 20% for pesticide expenses between "low-input" CMP and "high-yielding" CMP. 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. Finally, we also observe a parallel trend between the two groups implying that both groups 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. Yet, the size of the "low-input" cluster (more than 1/3 and up to 2/3 of the farmers) encourages us to think that this the latent class model with two group rather distinguish very high-yielding farmers from low-input and intermediate ones. 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. We tried to build a three-cluster latent class model to distinguish intermediate farmers from "true" low-input farmers. Yet, the results from the three-class model are less satisfactory. First, the three-class results lack from time consistency. Additionally, characteristics of low-input and intermediate class collapse between nitrogen and pesticide uses. The low-input class (from the yield point of view) has lower nitrogen expenses while higher pesticide expenses than the intermediate class. It might

confounded with CMP choices, especially as we do not give any structure to those CMP choices. Restricting ourself to this area for the exploratory analyses prevents us for such confusion. 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. Additionnally, to avoid the problem of missing data, we only consider farmers that were observed during the four considered years.

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. The constraints we introduced in our hidden Markov model permit to ensure that we distinguish for CMPs rather than for efficiency groups or other heterogeneity factors.

6 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 present the ex-post distribution of the random parameters from the CMP choice probability functions, *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 2.[21]

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 $\mu_{c1} \equiv 0$, $\forall c \in \mathcal{C}$. 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.

[21] Plotted distribution from these random parameters can be found in Appendix 7.3.

Table 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 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}.

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.[22] 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 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.

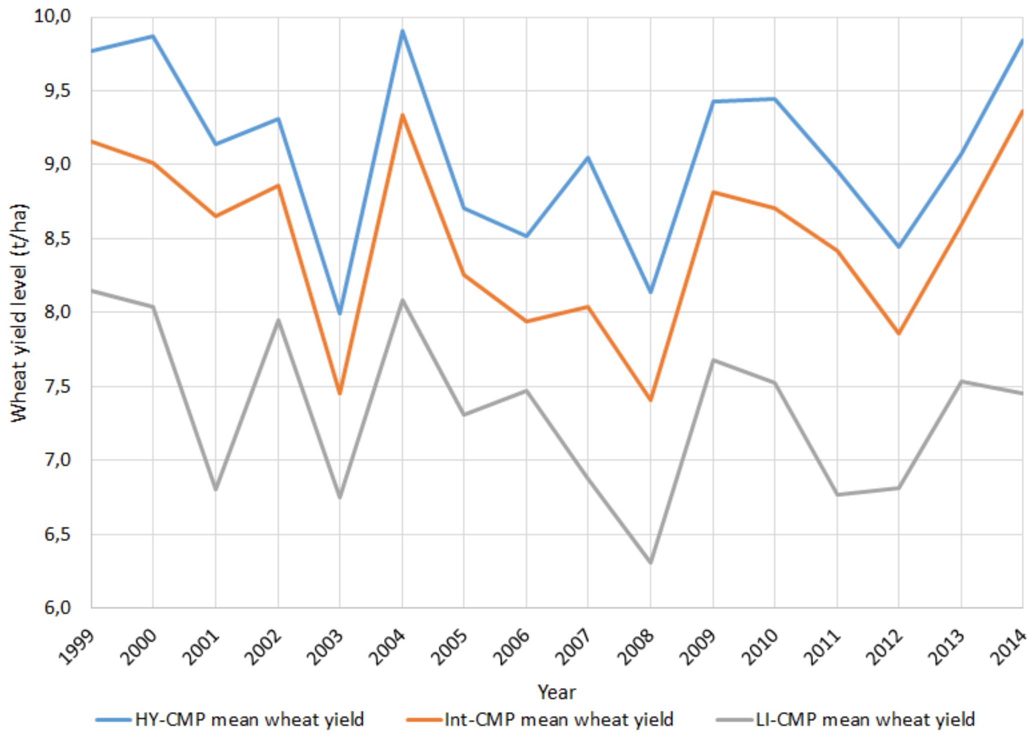
Let now consider the characteristics from the three categories that were distinguished thanks to the RPHMM: (*i*) high-input CMPs that are associated to larger levels of yield and input use (for both fertilizer and pesticide uses), (*ii*) intermediate CMPs with slightly lower yield and pesticide use levels but similar fertilizer use levels, (*iii*) low-input CMPs with lower yield and input use levels. Instead of presenting the mean and standard deviation from random ex-post distribution of parameters b^1 and a^c, we represent graphically the mean yield and input use levels of the three CMP categories.[23] 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 – respectively Figure 4.[24]

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

[23]Ex-post mean and standard deviation from random parameters b^1 and a^c can be found in Appendix 7.3 as well as the ex-post full distribution. Overall, these parameters benefit from a rather concentrated distribution which argues in favor of the stability of the built classes.

[24]Actually, it corresponds to $E[b_i^c] + d_t^0 + \delta_0^{\prime E[\mathbf{z}_{it}|r]}$.

Figure 3: Annual mean yields for the three CMP categories obtained with RPHMM, from 1999 to 2014



Source: Authors' calculations on CDER data.

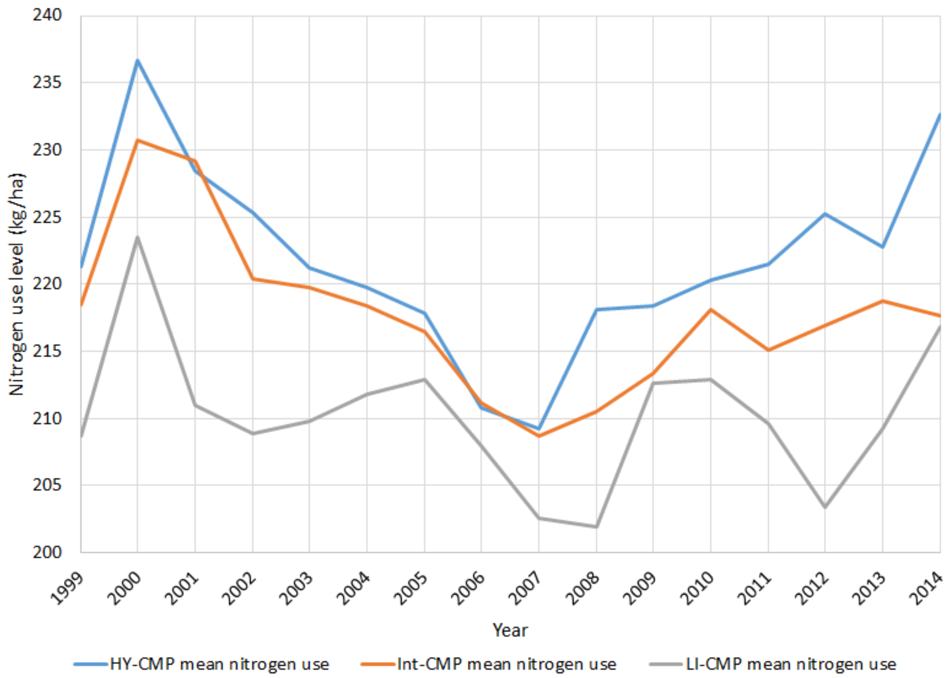
From these two figures, we can see that pesticide use and yield levels characterize the three CMP categories. On average, intermediate CMPs are using 1.8% less nitrogen and 16.5% less pesticide than high-input CMPs for a 6.7% decrease in yields. Low-input CMPs use on average 3.4% (respectively 5.2%) 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.

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%. Another element which is consistent with the agronomic view on CMPs is the fact that the most discriminating type of pesticides between CMP categories are fungicide and herbicide uses.[25] Indeed, 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.

[25]See Appendix 7.4 for the results from the detailed pesticide analysis.

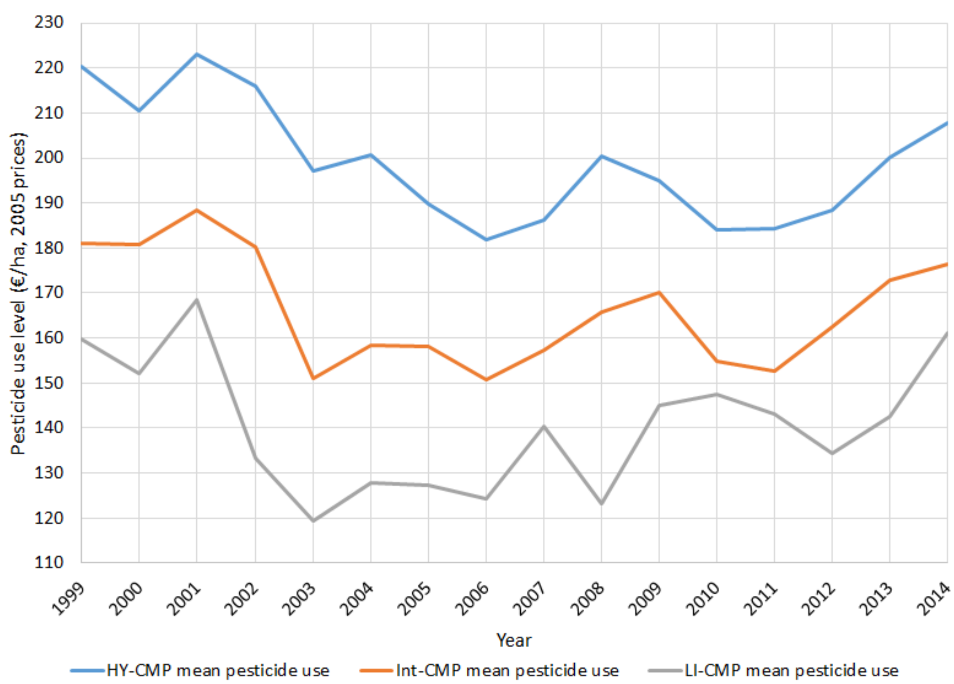
Figure 4: Annual mean input uses for the three CMP categories obtained with RPHMM, from 1999 to 2014

(a) Mean nitrogen use in each CMP type



Source: Authors' calculations on CDER data.

(b) Mean pesticide use in each CMP type

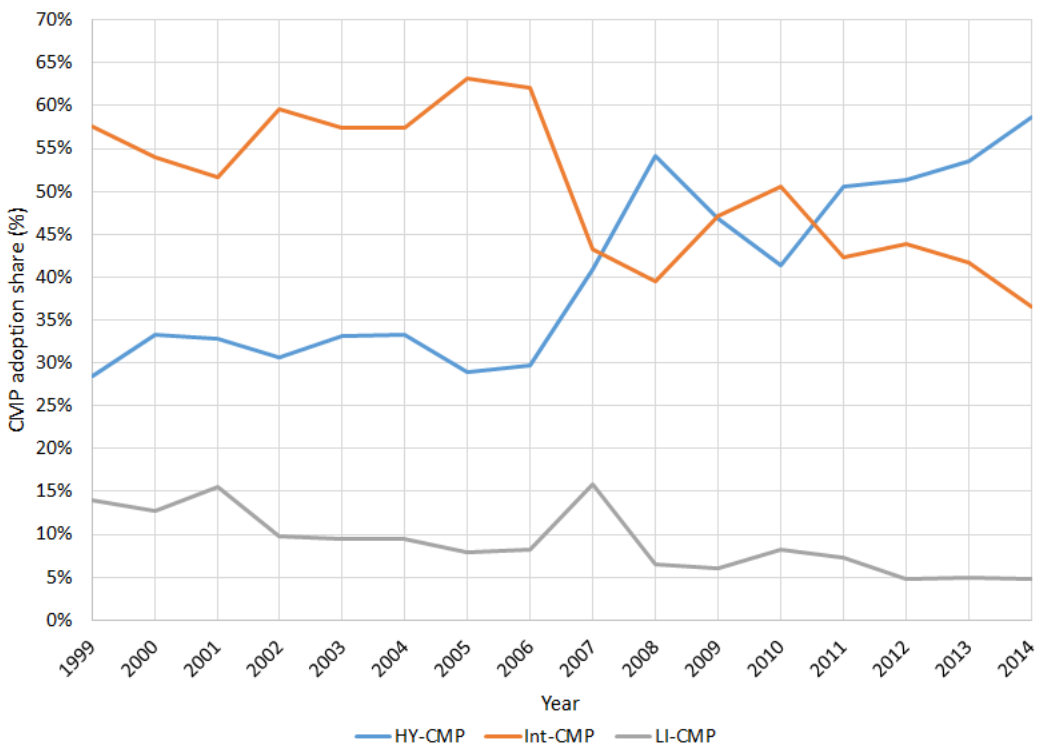


Source: Authors' calculations on CDER data.

On the contrary, pest infestations cannot be totally avoided, and insecticides remain the most efficient way to get rid of them. Thus, a less marked difference for insecticide uses compared to herbicide and fungicide uses.

Differences in input uses and in yields impact the expected profit associated to each CMP category. Figure 5 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 5: Estimated annual share of farmers who adopted a low-input CMP, from 1999 to 2014

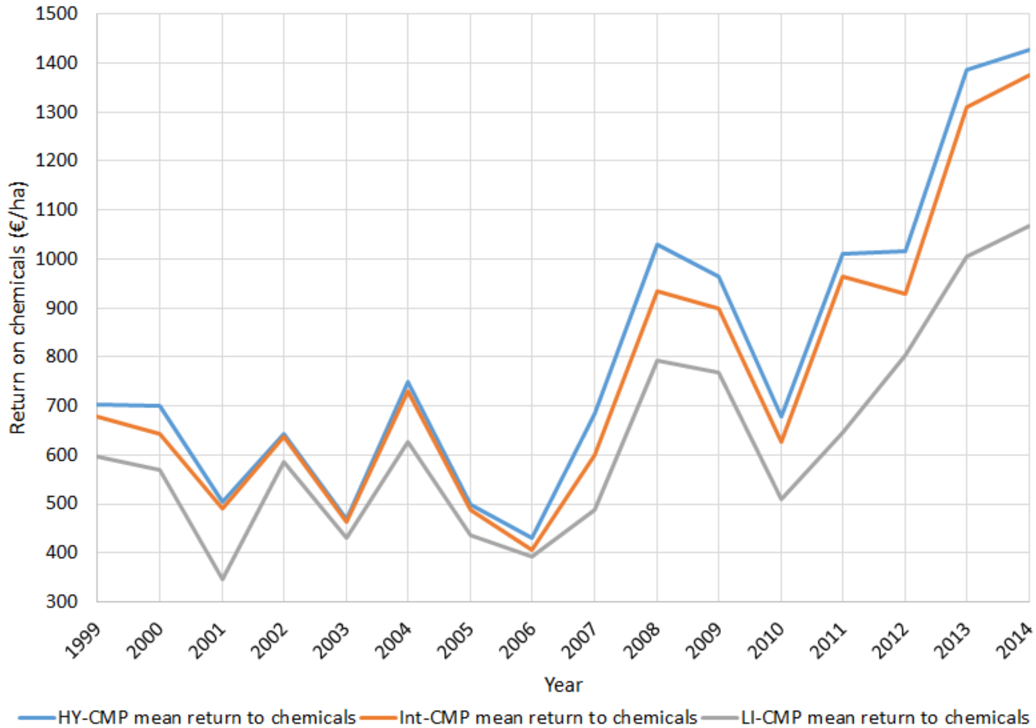


Source: Authors' calculations on CDER data.

Figure 6 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.

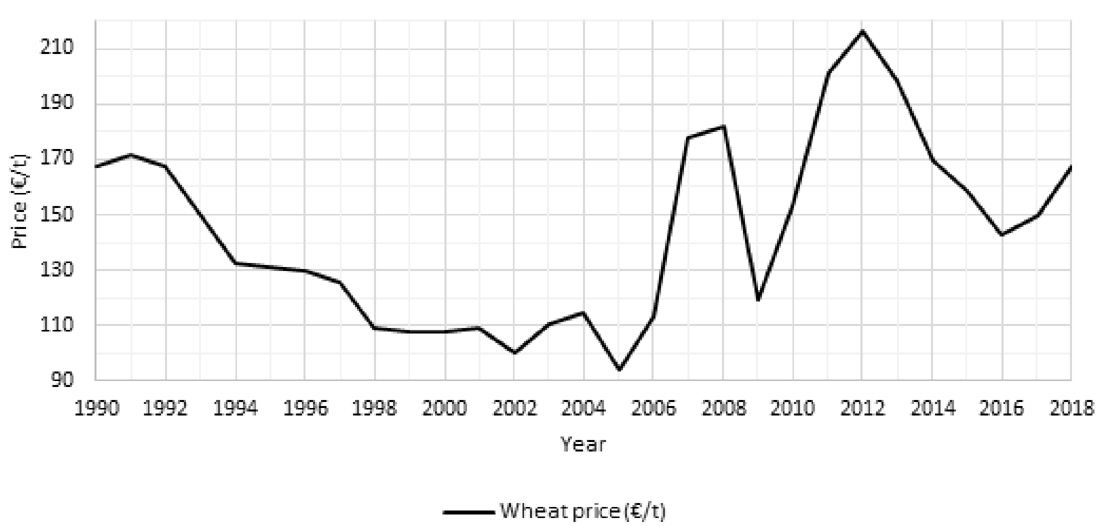
Figure 6: Estimated annual expected return for each CMP type (a) *versus* observed annual wheat prices (b)

(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.

Figure 6 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).

This idea seems to be confirmed by the results we get from the simulation work we produced to investigate into farmers' CMP choice reaction to price changes.[26] Overall, we found a limited responsiveness of CMP adoption on tax instrument. Even a 100% tax is insufficient to fill the gap between the revenue loss associated to lower yield levels associated to the low-input CMPs and their chemical input savings. On the other hand, a price premium instrument works better 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.

7 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'

[26] Detailed results from this simulation work can be found in Appendix 7.5.

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 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 that 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.[27] 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.

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 et al. (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.

[27] We investigated such extension. Yet, the introduction of such time trend in the initial probability of adoption entails convergence issues.

References

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.

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.

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.

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.

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.

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.

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.

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*.

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.

Eddy, S. R. (1996). “Hidden markov models”. In: *Current opinion in structural biology* 6.3, pp. 361–365.

Elmezain, M., A. Al-Hamadi, and B. Michaelis (2009). “Hand trajectory-based gesture spotting and recognition using HMM”. In: *2009 16th IEEE international conference on image processing (ICIP)*. IEEE, pp. 3577–3580.

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.

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.

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. 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.

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.

Goldfeld, S. M. and R. E. Quandt (1973). "A Markov model for switching regressions". In: *Journal of Econometrics* 1.1, pp. 3–15.

Greene, W. (2005). "Reconsidering heterogeneity in panel data estimators of the stochastic frontier model". In: *Journal of econometrics* 126.2, pp. 269–303.

Griliches, Z. (1957). "Hybrid corn: An exploration in the economics of technological change". In: *Econometrica, Journal of the Econometric Society*, pp. 501–522.

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.

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.

Henderson, J., S. Salzberg, and K. H. Fasman (1997). "Finding genes in DNA with a hidden Markov model". In: *Journal of Computational Biology* 4.2, pp. 127–141.

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.

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*.

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.

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.

— (July 2021). *Variable Input Allocation Among Crops: A Random Parameter Approach With Stochastic Constraints Using SAEM Algorithm.* Tech. rep. HAL.

Kuhn, E. and M. Lavielle (2005). “Maximum likelihood estimation in nonlinear mixed effects models”. In: *Computational statistics & data analysis* 49.4, pp. 1020–1038.

Larédo, P. and A. Hocdé (2014). *Variétés rustiques et itinéraires techniques économes en intrants.* Tech. rep. ASIRPA.

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.

Lintell Smith, G et al. (1992). *The effects of reduced nitrogen and weed competition on the yield of winter wheat.* Tech. rep.

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.

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.

Manogaran, G. et al. (2018). "Machine learning based big data processing framework for cancer diagnosis using hidden Markov model and GM clustering". In: *Wireless personal communications* 102.3, pp. 2099–2116.

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.

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.

Meynard, J.-M. (1985). "Construction d'itinéraires techniques pour la conduite du blé d'hiver". PhD thesis.

Meynard, J. (1991). "Pesticides et itinéraires techniques". In: *Phytosanitaires protection 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.

Monolix Methodology (2014). Version 4.3.3.

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.

Nguyen, N. T. et al. (2005). “Learning and detecting activities from movement trajectories using the hierarchical hidden Markov model”. In: *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR’05)*. Vol. 2. IEEE, pp. 955–960.

Orea, L. and S. C. Kumbhakar (2004). “Efficiency measurement using a latent class stochastic frontier model”. In: *Empirical economics* 29.1, pp. 169–183.

Poullette, B. (2018). “Enquête sur la structure des exploitations agricoles. En 2016, des exploitations moins nombreuses mais plus grandes”. In.

Prestat, E. et al. (2014). “FOAM (functional ontology assignments for metagenomes): a hidden Markov model (HMM) database with environmental focus”. In: *Nucleic acids research* 42.19, e145–e145.

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.

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.

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.

Rydén, T., T. Teräsvirta, and S. Åsbrink (1998). “Stylized facts of daily return series and the hidden Markov model”. In: *Journal of Applied Econometrics* 13.3, pp. 217–244.

Sexsmith, J. and U. Pittman (1963). “Effect of Nitrogen Fertilizers on Germination and Stand of Wild Oats”. In: *Weeds* 11.2, pp. 99–101.

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. (2011). “Selection and comparative advantage in technology adoption”. In: *Econometrica* 79.1, pp. 159–209.

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.

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.

Varga, A. and R. K. Moore (1990). “Hidden Markov model decomposition of speech and noise”. In: *International Conference on Acoustics, Speech, and Signal Processing*. IEEE, pp. 845–848.

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.

Welch, L. R. (2003). “Hidden Markov models and the Baum-Welch algorithm”. In: *IEEE Information Theory Society Newsletter* 53.4, pp. 10–13.

Appendices

7.1 Detailed estimation procedure

SAEM algorithm explicit forms

Terms $E^{(n)}[\ln \ell_i^C(\boldsymbol{\theta}|\mathbf{r}_{(i)}, \boldsymbol{\alpha}_i)|\boldsymbol{\zeta}_{(i)}]$ involve to compute expectations condition on the observed data $\boldsymbol{\zeta}_{(i)}$. 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 et al. (1999). It consists in divising 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 $\boldsymbol{\zeta}_{(i)}$ and the current value of parameter $\boldsymbol{\theta}$, $\boldsymbol{\theta}^{(n-1)}$:

1. **S-step**: Simulate $\left\{\hat{\boldsymbol{\alpha}}_i^{(n,m)} : m = 1, ..., R\right\}$ according to the conditional distribution $p\left(\boldsymbol{\alpha}_i \,|\, \boldsymbol{\zeta}_{(i)}, \boldsymbol{\theta}^{(n-1)}\right)$, for $i = 1, ..., N$

2. **E-step**: Given $\left\{\hat{\boldsymbol{\alpha}}_i^{(n,m)} : m = 1, ..., R\right\}$, evaluate the quantities $p_t(c \,|\, \mathbf{w}_{(i)}, \hat{\boldsymbol{\alpha}}_i^{(n,m)})$ and $s_t(c, d \,|\, \mathbf{w}_{(i)}, \hat{\boldsymbol{\alpha}}_i^{(n,m)})$, for $i = 1, ..., N$, $t = t_{(i)}, ..., T_{(i)}$, $c \in \mathcal{C}$ and $d \in \mathcal{C}$.

3. **SA-step**: update sufficient statistics according to

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

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

$$E^{(n)}[\ln L^C(\boldsymbol{\theta})|\boldsymbol{\zeta}]$$

$$\simeq$$

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

Then, we can realize the last step, *i.e.* the **M-step** that consists in updating parameter $\boldsymbol{\theta}$ according to:

$$\left(\begin{array}{ll} \boldsymbol{\alpha}_0{}^{(n)} & = N^{-1} \sum_{i=1}^N \mathbf{s}_i^{(\boldsymbol{\alpha},n)} \\ \boldsymbol{\Omega}^{(n)} & = N^{-1} \mathbf{s}_i^{(\boldsymbol{\alpha}\boldsymbol{\alpha},n)} - \boldsymbol{\alpha}_0{}^{(n)} (\boldsymbol{\alpha}_0{}^{(n)})' \\ \boldsymbol{\delta}^{(n)} & = \left(\sum_{i=1}^N \sum_{t=t_{(i)}}^{T_{(i)}} \sum_{c\in C} s_{it,c}^{(\zeta,n)} \tilde{\mathbf{z}}'_{it} (\boldsymbol{\Sigma}^{c(n-1)})^{-1} \tilde{\mathbf{z}}_{it} \right)^{-1} \\ & \quad \times \sum_{i=1}^N \sum_{t=t_{(i)}}^{T_{(i)}} \sum_{c\in C} \tilde{\mathbf{z}}'_{it} (\boldsymbol{\Sigma}^{c(n-1)})^{-1} \mathbf{s}_{it,c}^{(\boldsymbol{\varepsilon},n)} \\ \boldsymbol{\Sigma}^{c(n)} & = \left(\sum_{i=1}^N \sum_{t=t_{(i)}}^{T_{(i)}} s_{it,c}^{(\zeta,n)} \right)^{-1} \left(\begin{array}{c} \mathbf{s}_c^{(\boldsymbol{\varepsilon}\boldsymbol{\varepsilon},n)} - \sum_{i=1}^N \sum_{t=t_{(i)}}^{T_{(i)}} \tilde{\mathbf{z}}_{it} \boldsymbol{\delta}^{(n)} \left(\mathbf{s}_{it,c}^{(\boldsymbol{\varepsilon},n)} \right)' \\ - \sum_{i=1}^N \sum_{t=t_{(i)}}^{T_{(i)}} \left(\mathbf{z}_{it} \boldsymbol{\delta}^{(n)} \left(\mathbf{s}_{it,c}^{(\boldsymbol{\varepsilon},n)} \right)' \right)' \\ + \sum_{i=1}^N \sum_{t=1}^T s_{it,c}^{(\zeta,n)} \tilde{\mathbf{z}}_{it} \boldsymbol{\delta}^{(n)} (\tilde{\mathbf{z}}_{it} \boldsymbol{\delta}^{(n)})' \end{array} \right) \end{array} \right).$$

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:

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

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 $p(\boldsymbol{\alpha}|\boldsymbol{\zeta}_{(i)}, \boldsymbol{\theta}^{(n-1)})$ as stationary distribution, and we retain R MCMC draws for each i, $i = 1, ..., N$.[28] 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 $\hat{\boldsymbol{\alpha}}_{i,m}$ such that $\hat{\boldsymbol{\alpha}}_{i,m} \sim N(\hat{\boldsymbol{\alpha}}_{i,m-1}, \boldsymbol{\Psi})$ for $1 \leq i \leq N$ and $1 \leq m \leq R + R_{\text{burn}}$. We defined the acceptance rate as:

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

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

Stopping rule of the algorithm and diagnostic plots

As in Koutchadé et al. (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 et al., 2001). If the condition

$$\max_j \left(\frac{|\theta_{j,n} - \theta_{j,n-1}|}{|\theta_{j,n}| + \sigma_1}\right) < \sigma_2, \tag{15}$$

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 $\boldsymbol{\theta}_n$ achieved, at least approximately, the maximum of the considered likelihood function when the condition (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 $\boldsymbol{\theta}$ (Gu and Zhu, 2001). Third, we check graphically the convergence of parameters by plotting their values along iterations.

[28] In our case, we set $R = 1$ as we have many individuals.

Estimation of the variance of the estimates

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

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

and the variance of estimates by:

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

Estimation of the likelihood and model selection

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

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

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

$p(\mathbf{q}_{(i)}; \hat{\boldsymbol{\theta}})$ has no closed form, so we use the importance-sampling approach to estimate it. From prior distribution $\varphi(\boldsymbol{\alpha}_i - \hat{\boldsymbol{\alpha}}; \hat{\boldsymbol{\Omega}})$ as importance density, we draw independence sequence $\{\boldsymbol{\alpha}_{i,r} : r = 1, ..., R\}$ 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.[29] This estimator is unbiased and consistent as it variance decreases as $1/R$.

[29] We present the Forward-Backward algorithm at the end of this Appendix.

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 parameter to be estimated and N the number of observations.

Forward-Backward algorithm

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

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

for $i = 1, ..., N$, $t = t_{(i)}, ..., T_{(i)}$ 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 $(\mathbf{q}_{it_{(i)}}, ..., q_{it})$ given the random parameter $\boldsymbol{\alpha}_i$.

We can show that (see Maruotti, 2011):

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

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

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

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

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

for $i = 1, ..., N$, $t = t_{(i)}, ..., T_{(i)}$ and $c \in \mathcal{C}$. $\beta_{it}(c, \boldsymbol{\alpha}_i)$ denotes the probability of the partial sequence

$(\mathbf{q}_{it+1}, ..., \mathbf{q}_{iT_{(i)}})$ given that farmer i chooses CMP c at time t and given the random parameter $\boldsymbol{\alpha}_i$. We can compute this term iteratively by:

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

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

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

We also show that:

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

Estimation of the individual parameters and sequences of CMP

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

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

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

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

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

Yet, to compute this equation, we need to use the Viterbi algorithm (Rabiner, 1989).[30]

Viterbi algorithm

Let start by defining the following Viterbi path probability:

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

This term can be computed iteratively by:

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

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

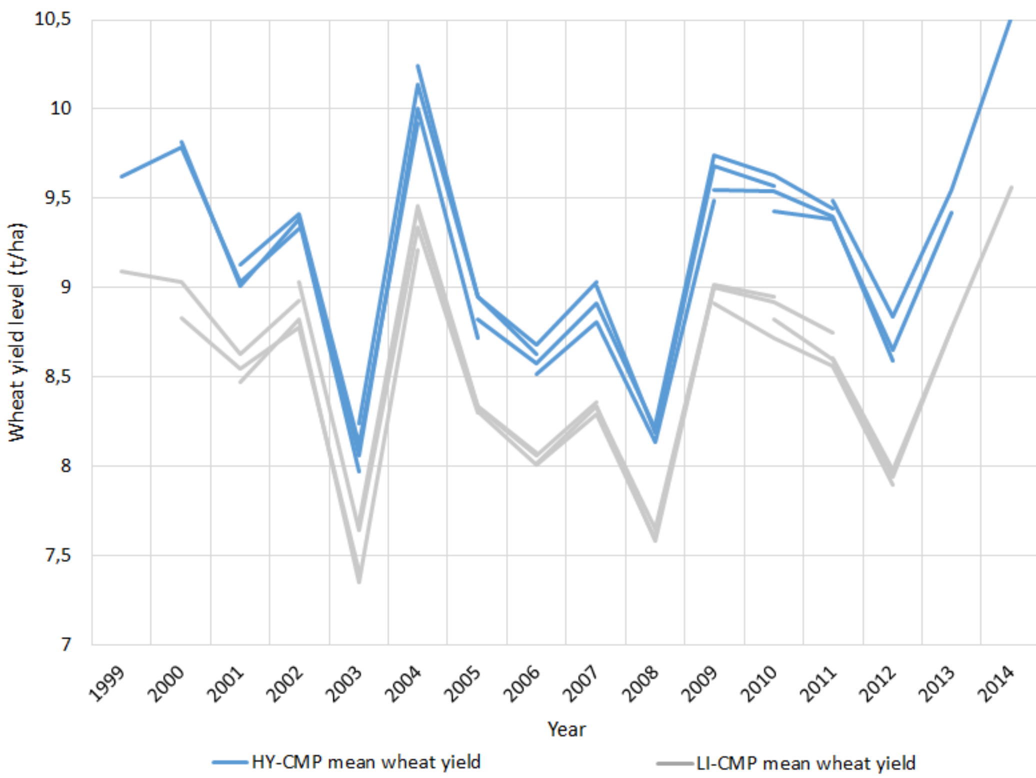
$$\begin{cases} r_{iT_{(i)}}^* & = \arg\max_{c \in C} v_{iT(i)}(c, \boldsymbol{\alpha}_i) \\ r_{it}^* & = \arg\max_{c \in C} v_{it+1}(c, \boldsymbol{\alpha}_i) p_{t+1}(r_{it+1}^*|c, \boldsymbol{\alpha}_i, \mathbf{w}_{t+1}) \end{cases}.$$

[30] As for the Forward-Backward algorithm, it is presented below.

7.2 Results from the "exploratory" clustering analyses

Figures 7 and 8 display the characteristics of the two classes we obtained with the latent class model while figure 9 shows the size of the "low-input" class.

Figure 7: Estimated annual mean yield of high-yielding and low-input farmers, from 1999 to 2014



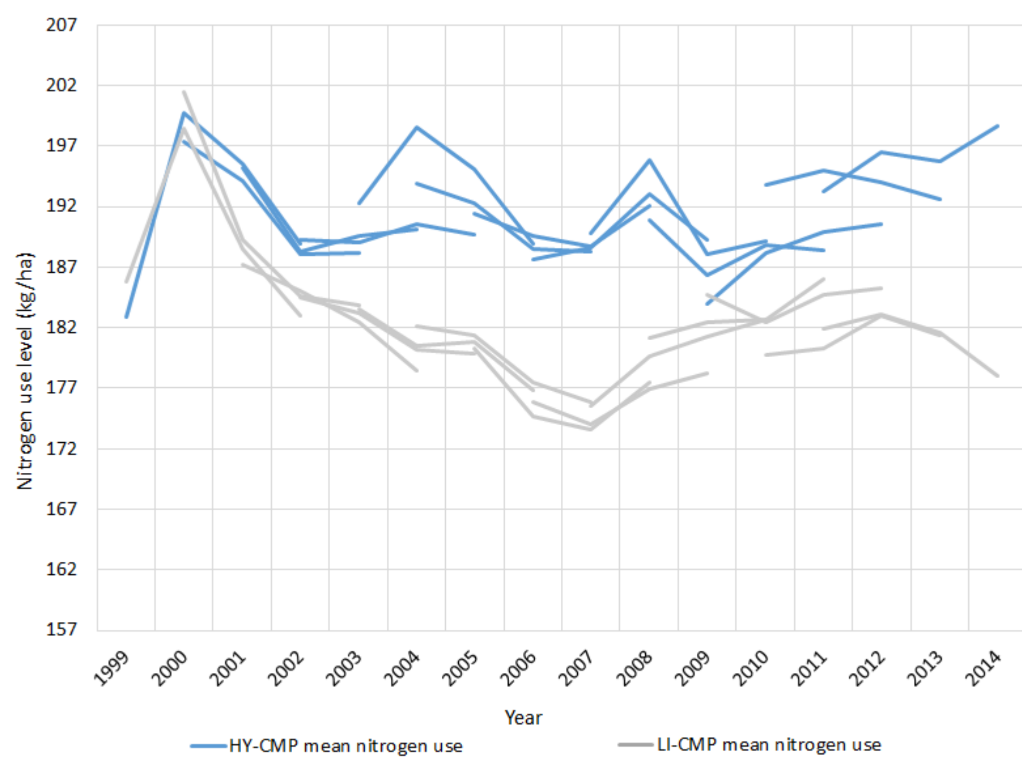
Source: Authors' calculations on CDER data.

Figure 8: 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

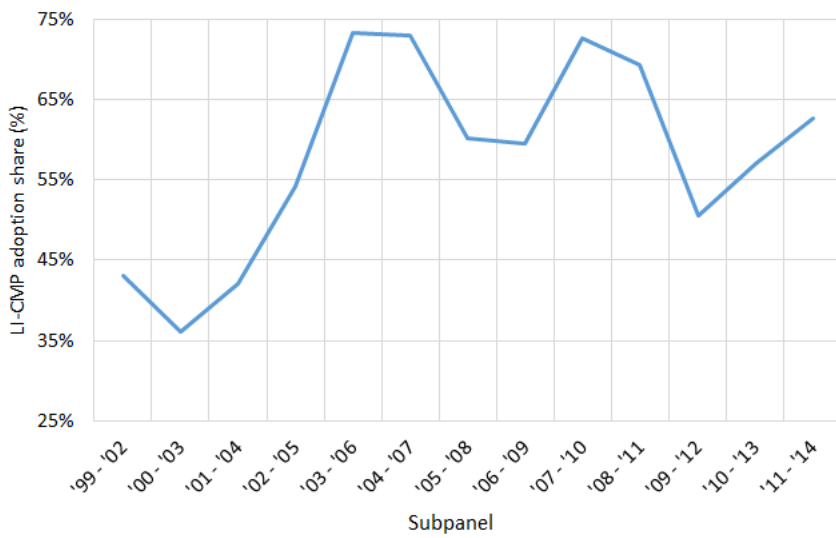
Source: Authors' calculations on CDER data.

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



Source: Authors' calculations on CDER data.

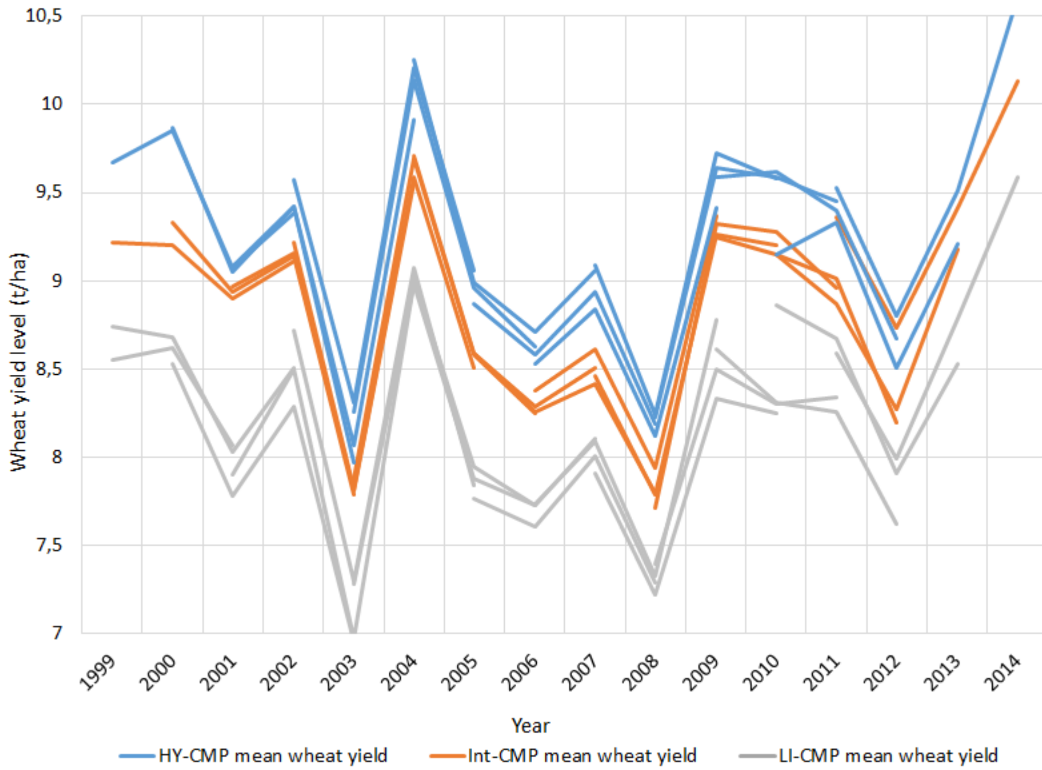
Figure 9: Estimated share of farmers who adopted a low-input CMP, from 1999 to 2014



Source: Authors' calculations on CDER data.

Figures 10 and 11 display the characteristics of the three classes we obtained with the latent class model.

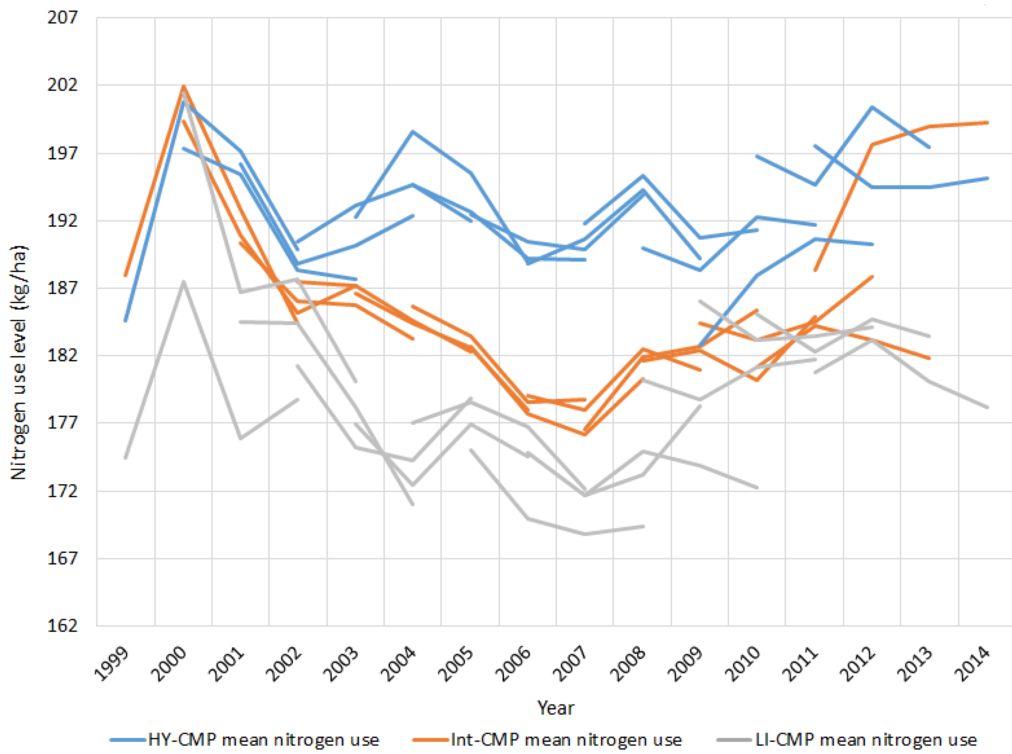
Figure 10: Estimated annual mean yields of high-yielding, intermediate and low-input CMPs, from 1999 to 2014



Source: Authors' calculations on CDER data.

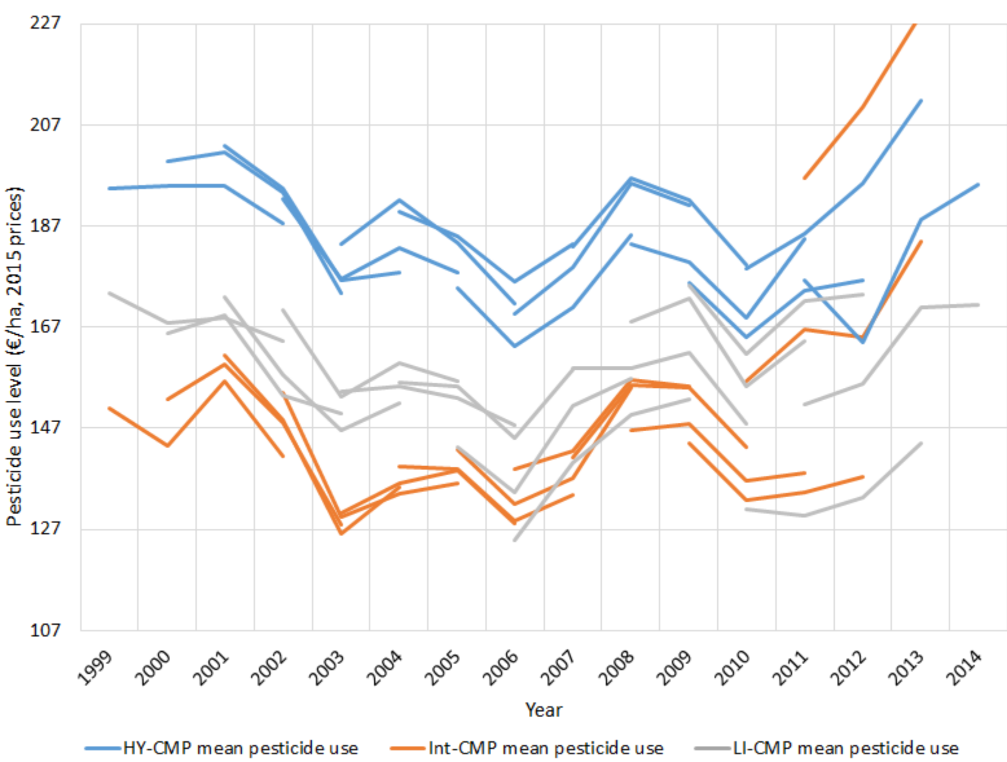
Figure 11: Estimated annual mean input uses of high-yielding, intermediate and low-input CMPs, from 1999 to 2014

(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

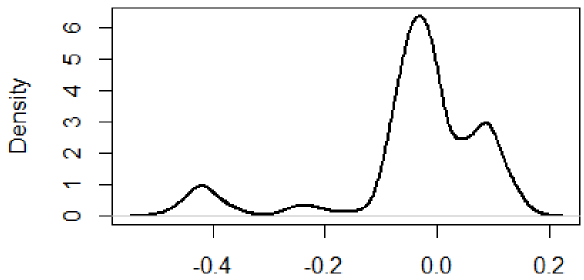


Source: Authors' calculations on CDER data.

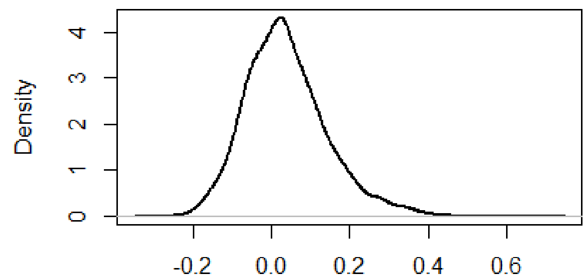
7.3 Distribution of random parameters from RPHMM

Figure 12: Distribution of cost parameters from the technology choice models

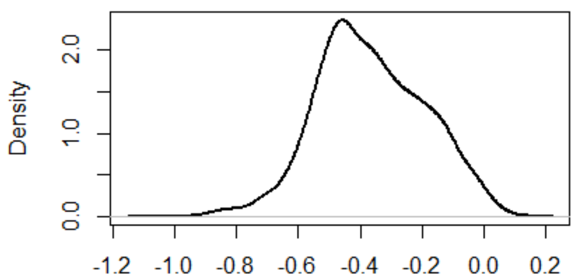
(a) Distribution of η_2



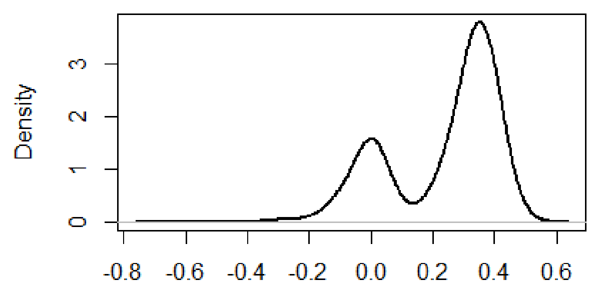
(b) Distribution of η_3



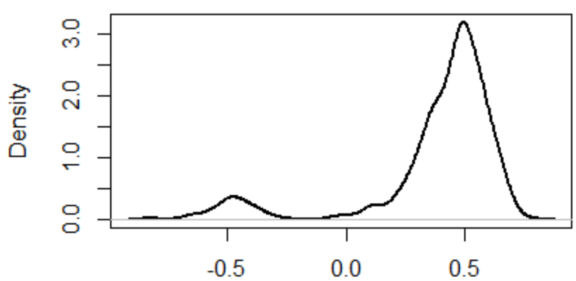
(c) Distribution of μ_{12}



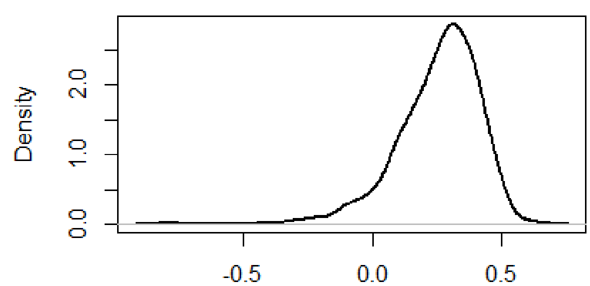
(d) Distribution of μ_{13}



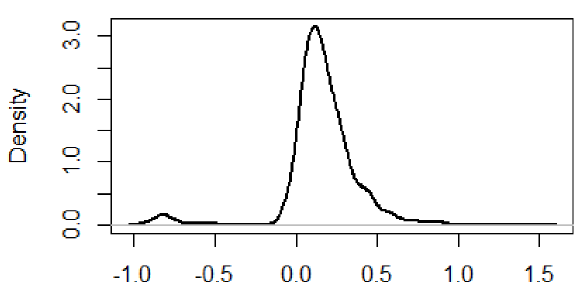
(e) Distribution of μ_{22}



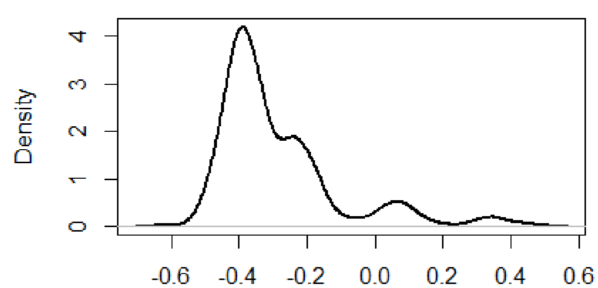
(f) Distribution of μ_{23}



(g) Distribution of μ_{32}



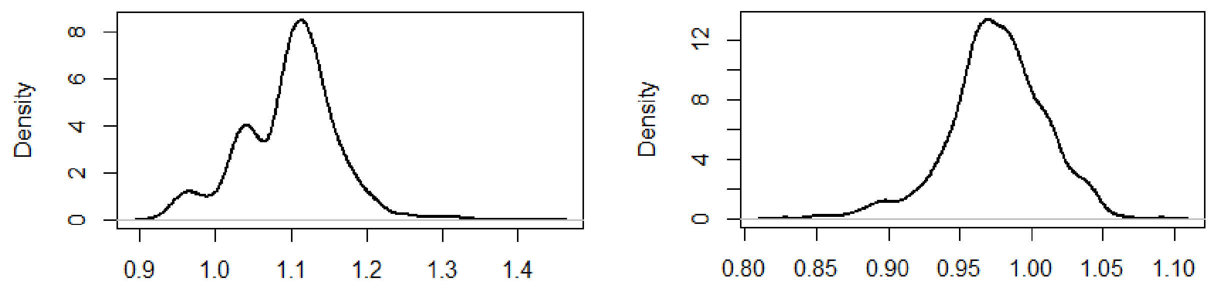
(h) Distribution of μ_{33}



Source: Authors' calculations on CDER data.

Figure 13: Distribution of scale parameters from the technology choice models

(a) Distribution of ρ_0 (b) Distribution of ρ



Source: Authors' calculations on CDER data.

Table 3: Estimation standard errors, mean and standard deviation from the ex-post distribution of random parameters from the yield and input use models

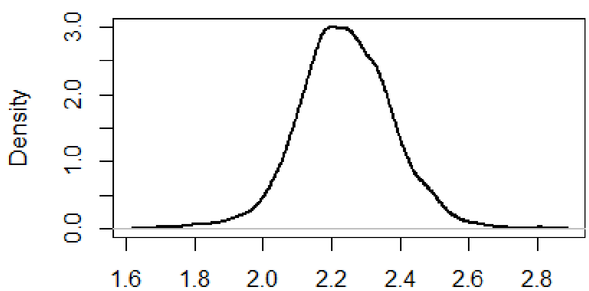
	Estimation se	Mean	sd
Input use & output levels			
b^1_{fert}	0.006	2.240	0.135
b^1_{pest}	0.002	1.957	0.299
b^1_{yield}	0.004	7.973	0.799
Discount parameters			
a^2_{fert}	0.003	0.937	0.012
a^2_{pest}	0.006	0.906	0.021
a^2_{yield}	0.005	0.952	0.009
a^3_{fert}	0.004	0.929	0.008
a^3_{pest}	0.005	0.929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

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

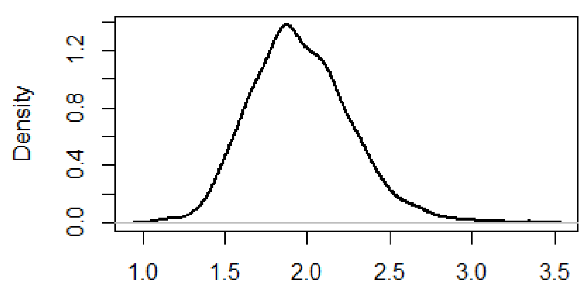
Source: Authors' calculations on CDER data.

Figure 14: Distribution of the input use and output levels parameters

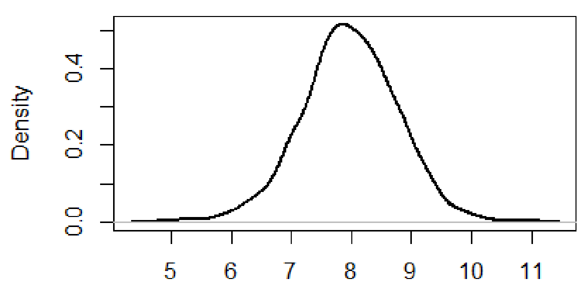
(a) Distribution of b^1_{fert}



(b) Distribution of b^1_{pest}

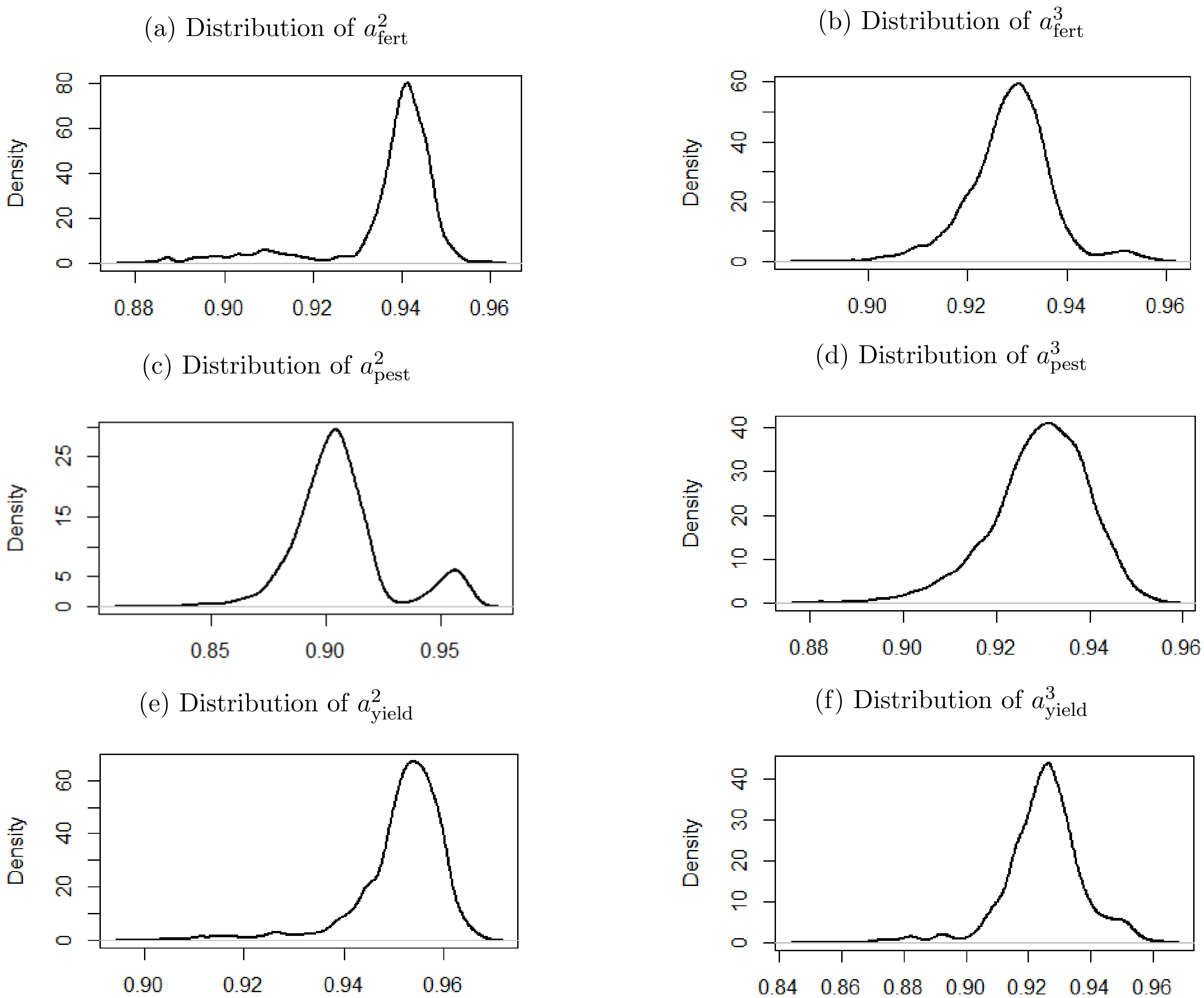


(c) Distribution of b^1_{yield}



Source: Authors' calculations on CDER data.

Figure 15: Distribution of discount parameters from the input use and output level models



(a) Distribution of a^2_{fert}

(b) Distribution of a^3_{fert}

(c) Distribution of a^2_{pest}

(d) Distribution of a^3_{pest}

(e) Distribution of a^2_{yield}

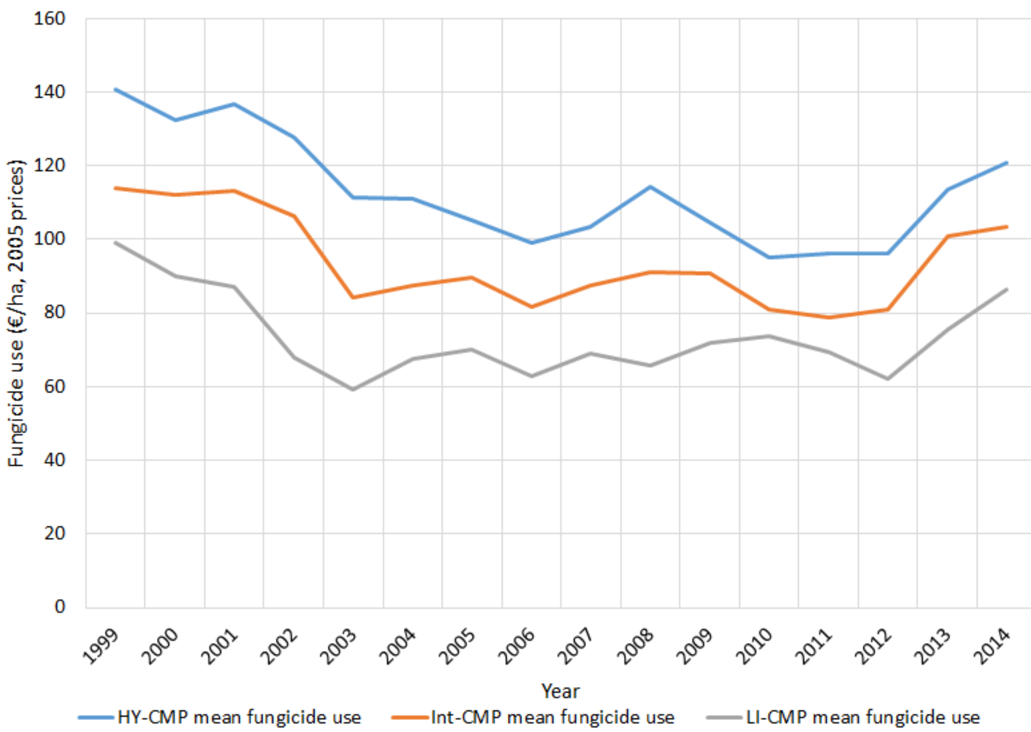
(f) Distribution of a^3_{yield}

Source: Authors' calculations on CDER data.

7.4 RPHMM results on detailed pesticide uses across CMP categories

From Figures 16 and 17, 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.

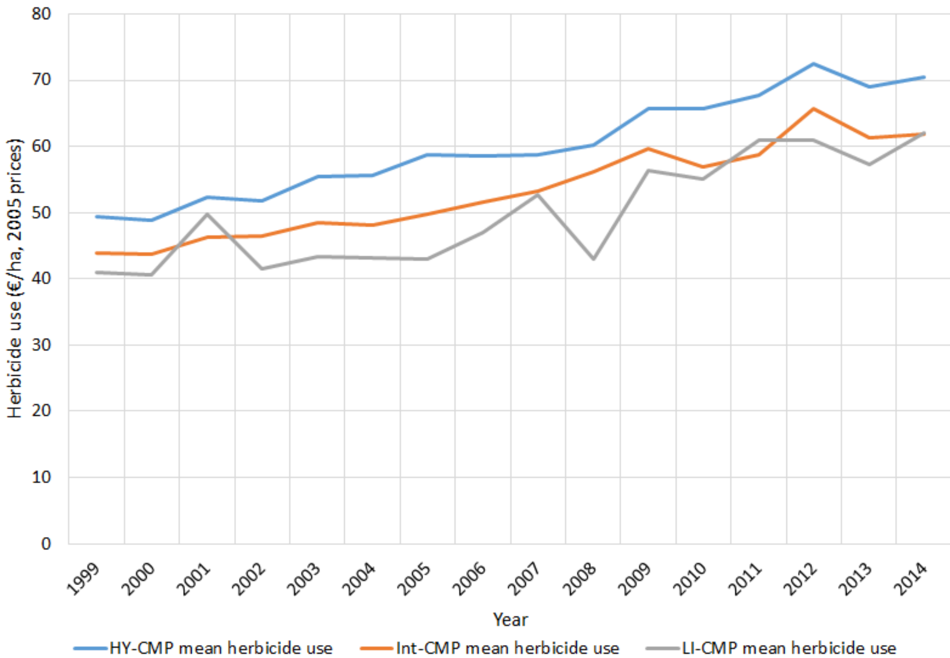
Figure 16: Annual mean fungicide uses for the three CMP categories obtained with RPHMM, from 1999 to 2014



Source: Authors' calculations on CDER data.

Figure 17: 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.

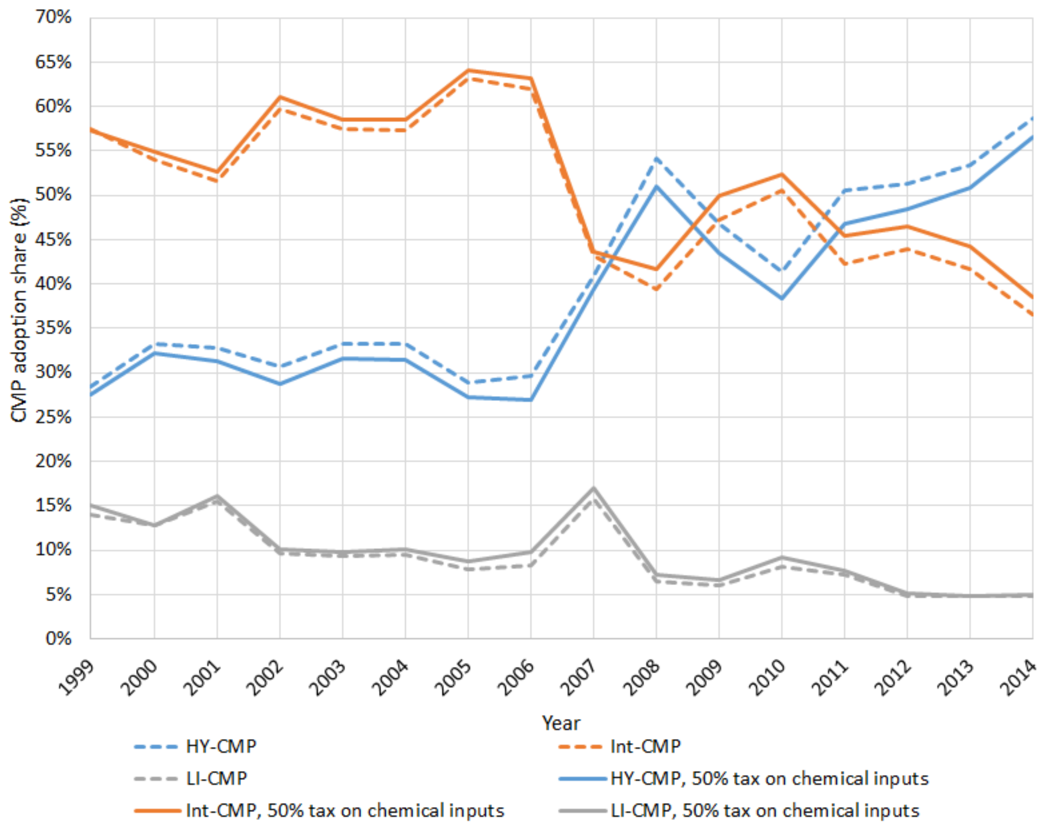
(b) Mean insecticide use in each CMP category

Source: Authors' calculations on CDER data.

7.5 Simulation results

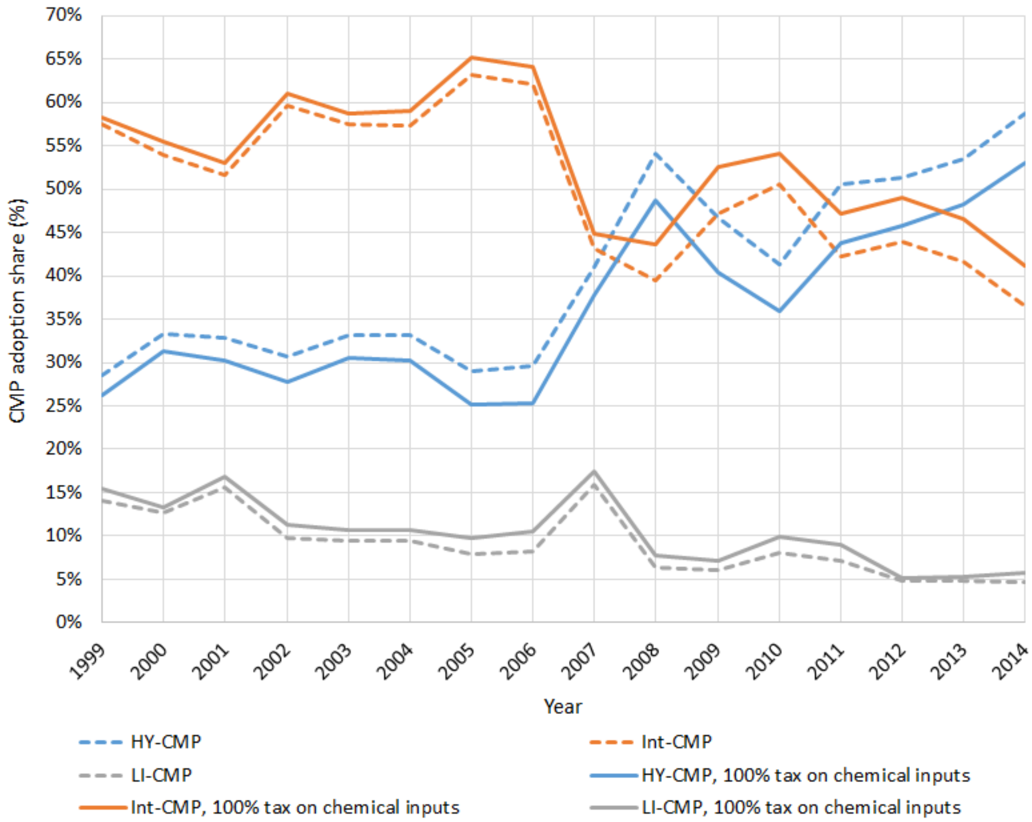
As evoked in the core of the article, we performed simulations based on the RPHMM results to investigate into farmers' CMP adoption decision. First, we simulate a 50% and 100% tax on chemical inputs. We present in Figure 18 – respectively in Figure 19 – the results from the 50% – respectively 100% – tax on chemical inputs.

Figure 18: Annual change in the CMP adoption share after simulating a 50% tax on chemical inputs



Source: Authors' calculations on CDER data.

Figure 19: Annual change in the CMP adoption share after simulating a 100% tax on chemical inputs

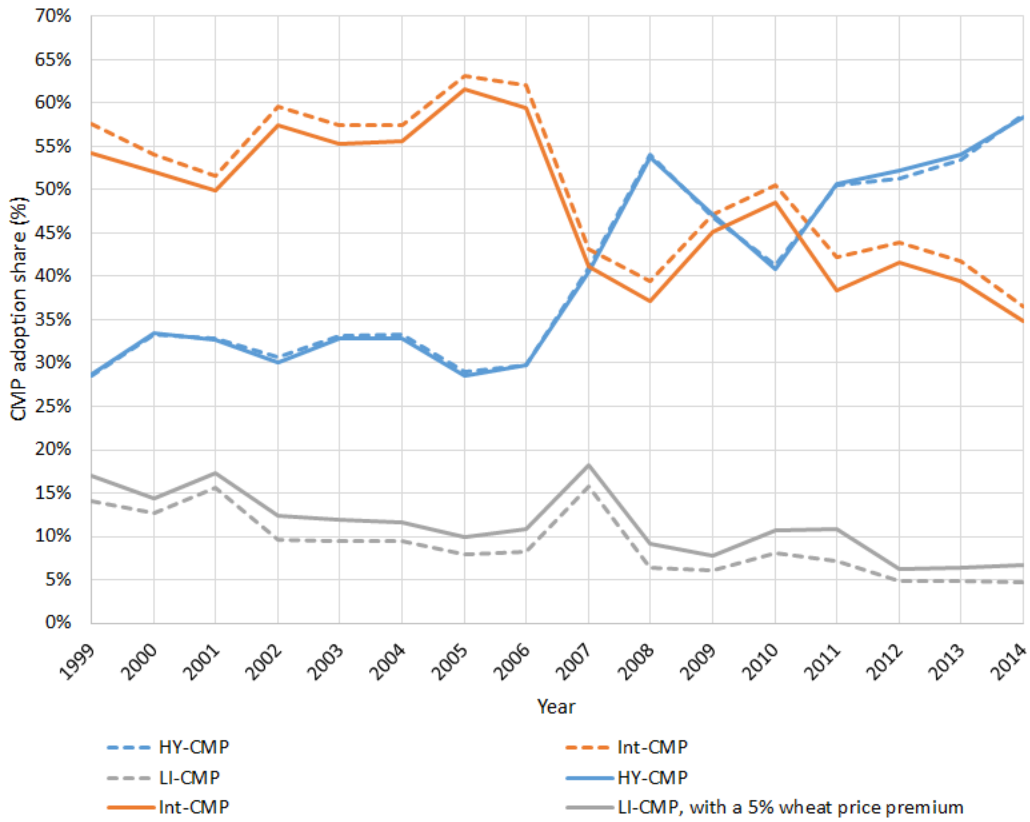


Source: Authors' calculations on CDER data.

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%, 10% and 20% simulations are presented in Figure 20 and 21.

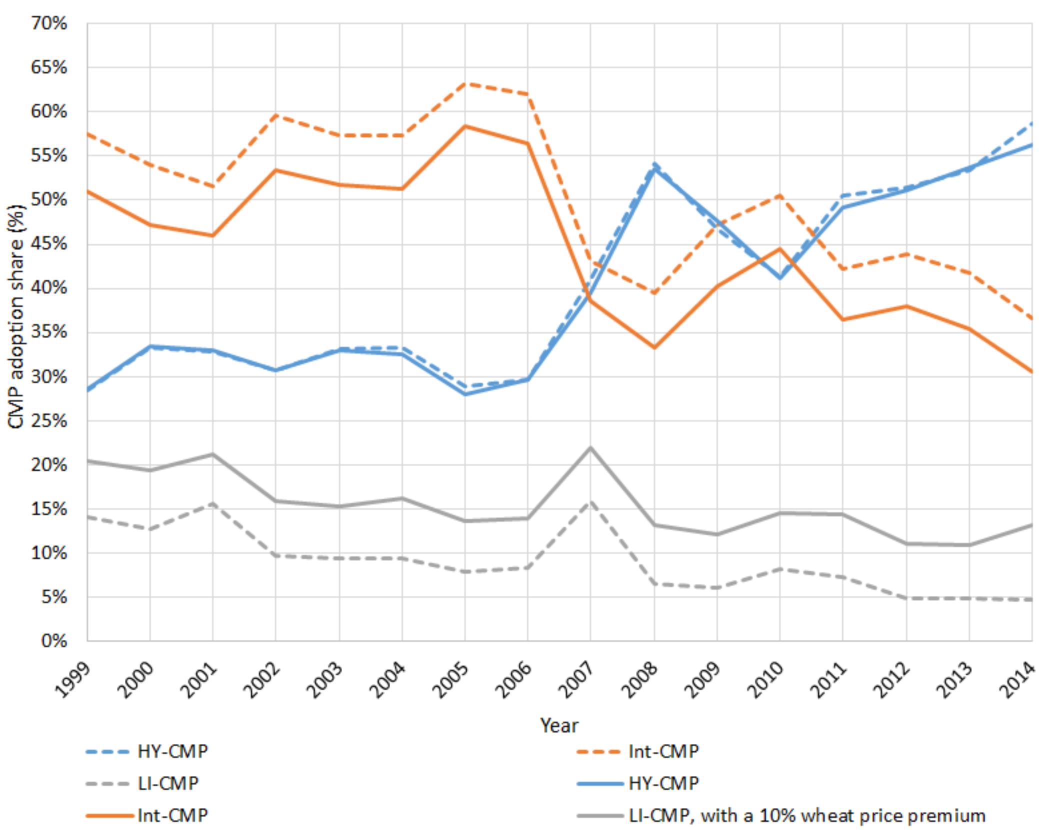
Figure 20: Annual change in the CMP adoption share after simulating price premiums for low-input wheat producers

(a) Results from the 5% price premium



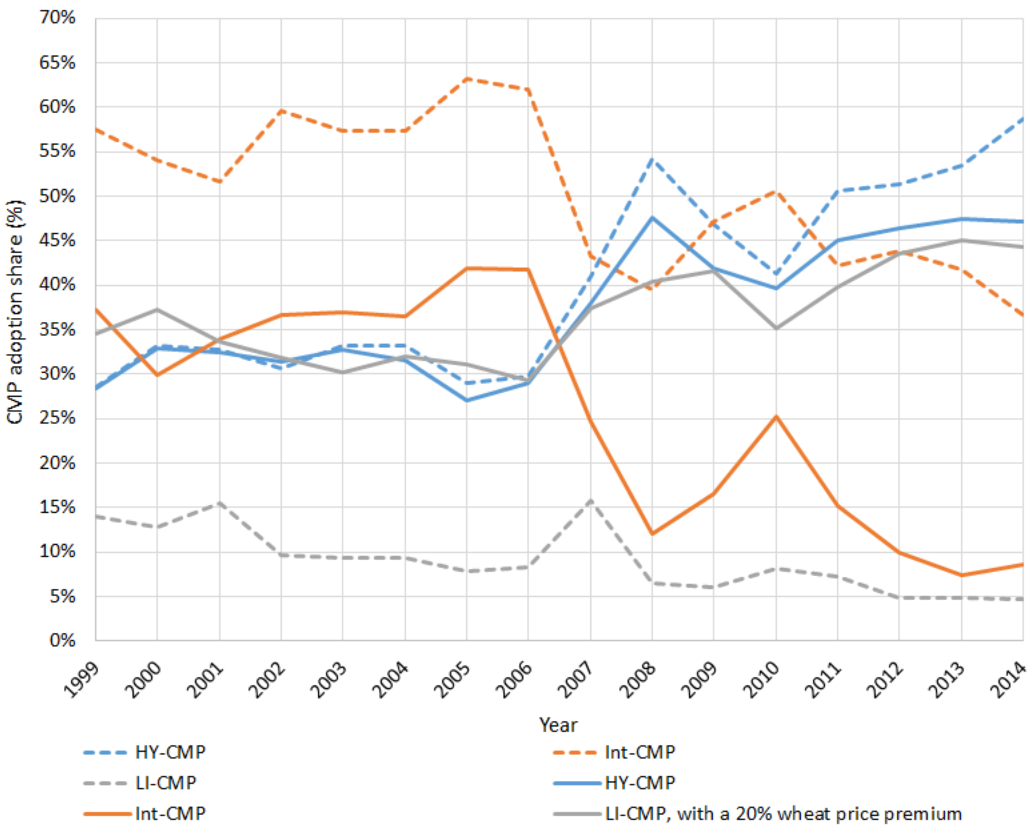
Source: Authors' calculations on CDER data.

(b) Result from the 10% price premium



Source: Authors' calculations on CDER data.

Figure 21: Annual change in the CMP adoption share after simulating a 20% price premiums for low-input wheat producers



Source: Authors' calculations on CDER data.