Accounting for Latent Cropping Management Practices Choices in Crop Production Models: a Random Parameter Hidden Markov Model approach

Esther Devilliers, Obafemi Philippe Koutchade, A. Carpentier

▶ **To cite this version:**

Esther Devilliers, Obafemi Philippe Koutchade, A. Carpentier. Accounting for Latent Cropping Management Practices Choices in Crop Production Models: a Random Parameter Hidden Markov Model approach. 2023. hal-04157564

HAL Id: hal-04157564
https://hal.science/hal-04157564

Preprint submitted on 10 Jul 2023

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

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

Accounting for Latent Cropping Management Practices Choices in Crop Production Models: a Random Parameter Hidden Markov Model approach

Esther Devilliers[1], Obafèmi Philippe Koutchadé[2], Alain Carpentier[2]

[1] Université de Lorraine, Université de Strasbourg, AgroParisTech, CNRS, INRAE, BETA, 54000, Nancy, France

[2] SMART-LERECO, INRAE, Rennes, France

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 heterogeneous technology. CMPs being latent in most economists' datasets and CMP changes entailing adjustment costs, 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. An illustration on French winter wheat producers of *La Marne* area uncovers very high-yielding, high-yielding and low-input CMPs. The characteristics of the low-input CMPs we uncover are very close to those tested by agronomists in the area covered by our data. We also 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 and thus to encourage farmers to adopt less intensive practices.

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. While economists generally advocate for implementing taxes to internalize the negative external effects of input uses, policymakers are reluctant to use taxes owing to their potential impact on farmers' income. Moreover, 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 (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 (dual) production choice models (Böcker and Finger, 2017). These estimates rely on limited input price variation and ignore that farmers can switch from pesticide dependent to pesticide saving crop production technologies, thereby widening the scope of pesticide use responses to changes in pesticide prices. 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 in crop production choices is likely to be useful for revisiting pesticide uses of farmers.

Agronomists' CMP concept is closely related to economists' production technology concept. CMP is defined as an ordered sequence of operations aimed to produce a given crop. For a given crop, CMPs differ according to their target yield levels and to their reliance on specific categories of inputs. A single production function cannot account for the variety of CMP

responses to input uses: different CMPs need to be considered as different crop production technology (Femenia and Letort, 2016). Unfortunately, targeted yield levels and many cropping techniques characterizing CMPs (*e.g.*, sowing dates, sowing densities, seed varieties) are unobserved in datasets used by agricultural production economists. 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. This across farm and year heterogeneity has significant effects on crop production and on farmers' choices even within areas of limited size (Koutchadé et al., 2021; 2018).

The micro-econometric model we propose in this article considers that the observed yields and chemical input use levels stem from a set of unobserved CMPs. We consider a hidden Markov model featuring farm specific random parameters and a 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) We assume that farmers choose their CMP at the beginning of the cropping season, thereby assuming that CMP choices are short run ones.[1] (b) CMP changes may entail adjustment costs, such as learning costs when farmers consider CMPs they are not familiar with. Thus, farmers tend to keep on using the same CMP unless sufficient changes in the regulatory or economic environments lead them to reconsider their CMP. Accordingly, we model farmers' CMP choice as a Markovian process. (c) 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 hidden Markov model by considering CMP choices as expected return maximizing discrete choices involving unobserved switching costs.

[1] CMP choice involves neither large investment costs nor significant inter-temporal trade-offs. Indeed, it shares much more similarities with standard crop variety choices (Michler et al., 2019; Suri, 2011) than with irrigation technology choices (Genius et al., 2014).

The random parameter hidden Markov (RPHM) model proposed in this article enables us to obtain original results on production practice choices in France based on a rich panel dataset of cost accounting observations. Our application considers wheat production in La Marne area, which is an administrative area (similar in size to a small US county) located in northeastern France. In particular, our estimation results reveal that the considered farmers mostly rely on relatively high-yielding CMPs, which are intensive in chemical inputs, for producing wheat. Yet, our results also reveal that a small fraction of farmers use pesticide saving CMPs. Importantly, our estimates tend to demonstrate that the CMP choices of the considered farmers display significant persistence over time but also respond to economic incentives, albeit to a limited extent.

Several other models feature components and properties of our RPHM model. 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 Cillero et al. (2019), Renner et al. (2021) or Dakpo et al. (2021). (b) Most of our model parameters are farm specific random parameters, as in the micro-econometric multi-crop models of Koutchadé et al. (2018, 2021) or in the input demand model of Los et al. (2021). (c) We assume that farmers' CMP choices can be modelled as Markovian processes. Standard latent class models can only describe two technology choice patterns (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 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.

The rest of the article is organized as follows. The second section discusses in more details the agronomic and economic principles of low-input CMPs. 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, fourth section gives an insight about the estimation procedure for our random parameter hidden Markov model. Fifth and sixth sections are dedicated respectively to the data presentation and our results description. Lastly, we provide some discussion elements and concluding remarks.

2. Background

2.1. Agronomic principles of low-input CMPs

We are interested in two specific CMP types: low-input CMPs and high-yielding CMPs. High-yielding CMPs are conceived to achieve high target yield levels by increasing seeding densities, choosing early seeding dates, relying on productive seed varieties and applying large amounts of, especially nitrogen, fertilizers. Early seeding dates tend to expose crops to pest outbreaks. Nitrogen fertilizer use tends to trigger competition by weeds. High seed densities, productive cultivars and high loads of nitrogen fertilizer tend to increase wheat susceptibility to diseases. Yet, availability of efficient chemical pesticides enables farmer to control the pest and weed pressures triggered by these high-yielding techniques (Oerke, 2006; Popp, 2011).

The basic principle of the conception of low-input CMPs, as developed by INRAE (then INRA) in the mid-1980s, is to lower target yield levels in order to lower chemical input uses, in particular pesticides (Loyce et al., 2008, 2012; Meynard, 1985, 1991). Lowering target yield levels directly reduces crop nutrition needs and, thereby, nitrogen fertilization uses. Low-input CMPs lower crop protection needs and, thus, pesticide uses by avoiding cropping techniques that increase pest and weed pressures. Also, hardy wheat cultivars – slightly less productive than very high-yielding ones but resistant to multiple disease – are complementary to the agronomic principles underlying the design of low-input 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 high-yielding CMPs.

2.2. Economic principles of low-input CMPs

Due to the relative scarcity of arable land in Western Europe, adopting high-yielding CMPs appeared to be the most profitable technological option for farmers, especially considering the price of pesticides which was relatively low, to benefit from high grain prices (Mahé and Rainelli, 1987; Meynard, 1991). Yet, the removal of the European Common Agricultural Policy price support in 1992 called into question the profitability of grain production from the late 1990s to the mid-2000s. High-yielding CMPs appeared to be much less profitable during this low grain price period.

Associated to stable prices for fungicides and insecticides, we expect the economic conditions from the late 1990s to 2006 to favor low-input CMPs adoption (Loyce et al., 2012; Meynard et al., 2009). If yield levels obtained with low-input CMPs are on average 10% lower than those obtained with high-yielding CMPs, farmers benefit from multiple input savings and thus cost reduction. Nitrogen fertilizer loads decrease by 10% from the high-

yielding CMPs to the low-input CMPs while the use of (mostly) fungicides and insecticides is reduced by around 30%. Plus, due to the lower sowing densities in low-input seed uses decrease by around 50% when using these CMPs. Finally, low-input CMPs are labor and fuel saving thanks to their lower expected pesticide application numbers.

2.3. CMPs and pesticide use

Pesticides are the keystone of conventional practices (Aubertot et al., 2007) whereas low-input CMPs were designed to allow a decrease in pesticide use (Loyce et al., 2012; Meynard et al., 2009). Thus, pesticide reduction policies should aim for a change in CMPs. Yet, most econometric studies that evaluate the pesticide demand elasticity assume – because they are generally based on short time dimension data – that the production technology is constant. Unfortunately, CMPs are rarely recorded in economic datasets. Yet, given the time period of the data we have at our disposal, the aforementioned economic characteristics of CMPs and INRA experimentation of low-input CMPs in France from the late 1990s, we expect some farmers have experimented low-input CMPs and to be able to recover them with appropriate statistical techniques.

To confirm this intuition, we performed a cluster analysis to uncover CMPs from farmers' yield and chemical input use levels. Cluster analysis was performed on 4-year overlapping subpanels.[2] Considering two classes yielded stable groups across the considered subpanels with characteristics that were coherent with the agronomic characteristics of low-input and high-yielding CMPs (an average 7% difference in yield levels for an average decrease of 5% and 20% in fertilizers and pesticide expenses). Yet, the size of the less intensive cluster is

[2] Standard cluster analysis assumes static clusters, which is problematic when considering panel data and short-run choice as CMP choice. In that case, either we can perform a cluster analysis for each year separately and suppose that the group belonging does vary across years or we can perform a cluster assuming that an individual belongs to the same group for the length of the considered observation period. Both approaches are unsatisfactory when considering technology adoption.

probably too large to correspond to a low-input cluster. Indeed, it represents more 1/3 and up to 2/3 of the farmers whereas low-input CMPs remains a rare technology in France (Femenia and Letort, 2016). Results from the three-class model lack from time consistency and characteristics of low-input and high-yielding class collapse between nitrogen and pesticide uses. It might indicate that we distinguish here for farmers' efficiency rather than production practices. Despite being mostly characterized by their target yield levels and their congruent chemical input use levels, CMPs cannot be identified by the direct comparison of yield and chemical input use levels across farms and years because of the numerous factors and events that can impact them.[3] Next section presents the structural equation modeling strategy we consider to uncover farmers' CMP choice and their associated production functions.

3. A random parameter hidden Markov model for production choices accounting for CMPs

3.1. Latent CMPs models

Our dataset being an unbalanced panel, let define $H_i = \{t(i), \dots, T(i)\}$ as the observation period of farmer i, where $i = 1, \dots, N$ and $H_i \subseteq \{1, \dots, T\}$. We assume that farmers can produce wheat by using a CMP among the C ones collected in set $\mathcal{C} = 1, \dots, C$. CMP indices, $c \in \mathcal{C}$, 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 $r_{it} \in \mathcal{C}$, is unobserved. Accordingly, variable r_{it} is considered

[3] 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 9.5t/ha of wheat in a good plot while it may only target 7.5t/ha 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.

as latent in our modelling framework.

Let variable y_{it}^{c} denote the wheat level obtained by farmer i considering that this farmer used CMP c in year t and let vector $\mathbf{x}_{it}^{c} = (x_{j,it}^{c}: j \in \mathcal{J})$ denote the corresponding input use levels. The model chosen for vector $\mathbf{q}_{it}^{c} = (y_{it}^{c}, \mathbf{x}_{it}^{c})$ is given by:

$$\begin{cases} y_{it}^{c} = b_{y,i}^{c} + d_{y,t,0} + \boldsymbol{\delta}_{y,0}'\mathbf{z}_{it} + \varepsilon_{y,it}^{c} \\ \boldsymbol{x}_{it}^{c} = \boldsymbol{b}_{x,i}^{c} + \boldsymbol{d}_{x,t,0} + \boldsymbol{\Delta}_{x,0}'\mathbf{z}_{it} + \boldsymbol{\varepsilon}_{x,it}^{c} \end{cases}, \text{for } c \in \mathcal{C} \qquad (1)$$

where vector $\mathbf{z}_{it}$ collects farm characteristics (*e.g.*, arable land area) and $\boldsymbol{\Delta}_0 = (\boldsymbol{\delta}_{y,0}, \boldsymbol{\Delta}_{x,0})$ their effects in the model of $\mathbf{q}_{it}^{c}$. We assume that $\mathbf{z}_{it}$ impacts 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 year specific fixed effects $\mathbf{d}_{t,0} = (d_{y,t,0}, \mathbf{d}_{x,t,0})$, depend neither on farms and nor on CMPs. Estimates terms $\mathbf{d}_{t,0}$ capture the effects on CMP specific yield and input use levels of price ratios, which vary much more across years than across farms, and of weather conditions, which are homogenous across farms in the considered sample. These homogeneity assumptions are admittedly restrictive. Identifying CMP specific effects from year specific factors is difficult in a latent CMP framework. In our model, the effects of CMPs and production conditions on wheat yields and input uses are captured by farm and CMP specific random parameters $\mathbf{b}_{i}^{c} = (b_{y,i}^{c}, \mathbf{b}_{x,i}^{c})$, where $\mathbf{b}_{x,i}^{c} = (b_{x,j,i}^{c}: j \in \mathcal{J})$. Term $b_{y,i}^{c}$ denotes the wheat target yield level of CMP c as this practice is implemented by farmer i. Similarly, term $b_{x,j,i}^{c}$ denotes the requirement in input j of CMP c. Error terms $\boldsymbol{\varepsilon}_{it}^{c} = (\varepsilon_{y,it}^{c}, \boldsymbol{\varepsilon}_{x,it}^{c})$ capture the effects of random events on wheat yield and input use levels that may depend on farms, years and CMPs. We assume that vectors $\mathbf{b}_{i}^{c}$, $\mathbf{z}_{it}$ and $\boldsymbol{\varepsilon}_{it}^{c}$ are mutually independent.[4] We also assume that error

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

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

We adopt a specific parameterization of random parameters $\mathbf{b}_{i}^{c}$ to ensure that our latent modelling framework empirically identifies CMPs. It defines terms $b_{y,i}^{c}$ and $b_{x,j,i}^{c}$ based on simple recursive schemes, $b_{y,i}^{c} = a_{y,i}^{c} b_{y,i}^{c-1}$ and $b_{x,j,i}^{c} = a_{x,j,i}^{c} b_{x,j,i}^{c-1}$, implying that $b_{y,i}^{c}$ and $b_{x,j,i}^{c}$ are given by the following simple formulae:

$$b_{y,i}^{c} = b_{y,i}^{1} \prod_{d=2}^{c} a_{y,i}^{d} \text{ , for } c \in \mathcal{C}_{a}, \tag{2}$$

and

$$b_{x,j,i}^{c} = b_{x,j,i}^{1} \prod_{d=2}^{c} a_{x,j,i}^{d} \text{ , for } c \in \mathcal{C}_{a} \text{ and } j \in \mathcal{J}, \tag{3}$$

where $\mathcal{C}_{a} = \{2, \dots, C\}$. The relative properties of high-yielding *versus* low-input CMPs and our defining c as an index that decreases with target yield level imply that expected yield and chemical input use levels $b_{y,i}^{c}$ and $b_{x,j,i}^{c}$ are non-negative and decrease in c. To ensure the identification of production technologies of distinct intensity levels we assume that $b_{y,i}^{1} > 0$, $a_{y,i}^{c} \in [0,1]$, $b_{x,j,i}^{1} > 0$ and $a_{x,j,i}^{c} \in [0,1]$, for $c \in \mathcal{C}_{a}$ and $j \in \mathcal{J}$. These conditions can be enforced by using suitable probability distributions for random parameter vectors $\boldsymbol{\gamma}_{i} = \left(\mathbf{b}_{i}^{1}, \mathbf{a}_{y,i}, \mathbf{a}_{x,i}\right)$, where $\mathbf{a}_{y,i} = (\mathbf{a}_{y,i}^{c} : c \in \mathcal{C}_{a})$, $\mathbf{a}_{x,i}^{c} = (a_{x,j,i}^{c} : j \in \mathcal{J})$ and $\mathbf{a}_{x,i} = (\mathbf{a}_{x,i}^{c} : c \in \mathcal{C}_{a})$.[5]

[5] 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 $i = 1, \dots, N$. The purposes of this "fixed parameter" specification were (*i*) to secure the identification of the model parameters and (*ii*) to facilitate the comparison of the yield levels and chemical input uses across CMPs. Surprisingly enough, estimating the model with fixed parameters $\mathbf{a}_{y,0}$ and $\mathbf{a}_{x,0}$ was impossible due to convergence issues while the estimated joint probability distribution of random parameters $\mathbf{a}_{y,i}$ and $\mathbf{a}_{x,i}$ display limited (although statistically significant) variability.

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:

$$f(\mathbf{q}_{it}|r_{it} = c, \boldsymbol{\gamma}_i, \mathbf{z}_{it}) = \varphi\left(\mathbf{q}_{it} - \mathbf{b}_i^c - \mathbf{d}_{t,0} - \boldsymbol{\Delta}_0 \mathbf{z}_{it} \,;\, \boldsymbol{\Sigma}_0^c\right), \tag{4}$$

where $\varphi(\mathbf{m}, \boldsymbol{\Xi})$ is the probability distribution function of $\mathcal{N}(\mathbf{0}, \boldsymbol{\Xi})$ at $\mathbf{m}$. The link between the observed input use and yield levels and the set of their latent CMP specific counterparts is given by following regime switching equation:

$$\mathbf{q}_{it} = \sum_{c \in \mathcal{C}} 1(r_{it} = c)\, \mathbf{q}_{it}^c = \sum_{c \in \mathcal{C}} r_{it}^c\, \mathbf{q}_{it}^c, \tag{5}$$

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. Defining the unconditional probability distribution of $\mathbf{q}_{it}$ requires specifying the probability distribution of $r_{it} = c$ as farmers' CMP choice is unobserved (and we already assume a probability distribution for $\boldsymbol{\gamma}_i$).

3.2. CMP choice model

The model of r_{it} assumes that farmers chose their CMP by maximizing their expected profit level up to CMP switch costs. Let $w_{y,t}$ denote wheat price paid to farmers and $w_{x,j,t}$ denote the price paid by farmers for input j in year t. The return to chemical inputs of wheat production obtained by farmer i in year t is given by $\tilde{\pi}_{it}^c = w_{y,t} y_{it}^c - \mathbf{w}_{x,t}' \mathbf{x}_{it}^c$, when this farmer uses CMP c. If input purchase prices $\mathbf{w}_{x,t} = (\mathbf{w}_{x,j,t} : j \in \mathcal{J})$ 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 denotes 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:

$$\pi_{it}^{c} = E\left[w_{y,t}\middle|\omega_{it}\right]b_{y,i}^{c} - \mathbf{w}_{x,t}'\mathbf{b}_{x,i}^{c} + \kappa_{it}, \quad (6)$$

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

Terms $E\left[\mathbf{d}_{x,t,0}\middle|\omega_{it}\right]$ and $E\left[d_{y,t,0}\middle|\omega_{it}\right]$ 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 modelling 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 $E\left[w_{y,t}\middle|\omega_{it}\right] = w_{y,t-1}$ (Koutchadé et al., 2018, 2021). It implies that $\pi_{it}^{c} = w_{y,t-1}b_{y,i}^{c} - \mathbf{w}_{x,t}'\mathbf{b}_{x,i}^{c} + \eta_{it}$.[6]

Even if we expect the relative profitability levels of the CMPs under consideration to significantly vary across years, we do not expect farmers to change their CMP frequently because of transition costs. Farmers are expected to tend to stick to the CMP they used the previous year, implying that CMP choice is a dynamic process. The sequence of CMP choices r_{it} is assumed to follow a (possibly) farmer specific first order Markov chain given expected crop returns $\boldsymbol{\pi}_{it} = (\pi_{it}^{c}: c \in C)$. We define the transition probabilities of the CMP choice process as functions of expected returns $\boldsymbol{\pi}_{it}$ and of implicit CMP switch costs. 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 $p(d|c, \boldsymbol{\alpha}_i, \mathbf{w}_t) = P[r_{it} = d\ |r_{it-1} = c, \mathbf{w}_t, \boldsymbol{\alpha}_i]$ denote probability of farmer i using CMP d in

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

year t conditionally on this farmer using CMP c in year $t-1$, on random terms $\boldsymbol{\alpha}_i$ and on expected price levels $\mathbf{w}_t = (w_{y,t-1}, \mathbf{w}_{x,t})$. Transition probability function $p(d|c, \boldsymbol{\alpha}_i, \mathbf{w}_t)$ is defined by:

$$p(d|c, \boldsymbol{\alpha}_i, \mathbf{w}_t) = \frac{\exp \rho_i(\pi_{it}^d - \mu_i^{d|c})}{\sum_{k\in\mathcal{C}} \exp \rho_i(\pi_{it}^k - \mu_i^{k|c})}, \text{ for } (c,d) \in \mathcal{C}\times\mathcal{C}, \tag{7}$$

where $\mu_i^{d|c}$ is the switching cost from CMP c to d.[7] The functional form of $p(d|c, \boldsymbol{\alpha}_i, \mathbf{w}_t)$ is that of a (mixed) Multinomial Logit model. This discrete choice model describes the choice of a CMP from set $\mathcal{C}$ in year t by farmer i assuming that this farmer used CMP c in year $t-1$.[8] The underlying discrete choice model is an expected return maximization problem given by $r_{it} = \arg\max_{k\in C}\{\pi_{it}^k - \mu_i^{k|c} + \rho_i^{-1} e_{it}^{k|c}\}$. The probability functions displayed in Equation (7) correspond to this model of r_{it} if elements of error term vector $\mathbf{e}_{it}^{|c} = \left(e_{it}^{d|c} \mid d \in \mathcal{C}\right)$, 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 $e_{it}^{d|c}$.

Equation (7) defines the transition probabilities of the CMP choice process. Determining the probability function of the (unobserved) sequence of CMP choices of farmer *i* requires to define the probability of this farmer choosing CMP *c* in year $t(i)$. Let $p_0\left(c \middle| \boldsymbol{\alpha}_i, \mathbf{w}_{t(i)}\right)$ denote such probability given $\boldsymbol{\alpha}_i$ and $\mathbf{w}_{t(i)}$. We assume that the functional form $p_0\left(c \middle| \boldsymbol{\alpha}_i, \mathbf{w}_{t(i)}\right)$ is given by:

$$p_0\left(c \middle| \boldsymbol{\alpha}_i, \mathbf{w}_{t(i)}\right) = \frac{\exp \sigma_i(\pi_{it(i)}^c - \eta_i^c)}{\sum_{k\in\mathcal{C}} \exp \sigma_i(\pi_{it(i)}^k - \eta_i^k)}, \text{ for } c \in \mathcal{C}. \tag{8}$$

[7] We impose the normalization constraints stating that $\mu_i^{1|c} = 0$ for $c \in \mathcal{C}$. We choose CMP 1, the most intensive CMP, as the benchmark choice because we expect most farmers to use high yielding CMPs.
[8] 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 $d \in \mathcal{C}$.

where term η_i^c captures the effects of farmer specific costs or motives that tend to deter farmer *i* from choosing CMP *c*.[9] This functional form is inspired by the Multinomial Logit probability function associated to the expected profit maximization problem given by $\max\{_{k\in\mathcal{C}}\, \pi_{it(i)}^k - \eta_i^k + \sigma_i^{-1} e_{it(i)}^k\}$. Term σ_i is a positive farm specific parameter scaling the effects of error terms $e_{it(i)}^k$.

Combining Equations (7) and (8), we can finally define the probability function of the sequence of CMP choices of the sampled farmers. Let $\mathbf{r}_i = (r_{it}: t \in H_i)$ define the sequence of latent CMP choices of farmer i. Let function $P(\mathbf{r}_i|\boldsymbol{\alpha}_i, \mathbf{w}_i)$ denote the probability function of $\mathbf{r}_i$ given $\boldsymbol{\alpha}_i$ and $\mathbf{w}_i = (\mathbf{w}_t: t \in H_i)$. Under our assumption set, the probability function $P_i(\mathbf{r}_i|\boldsymbol{\alpha}_i)$ of CMP choice sequence $\mathbf{r}_i$ given $\boldsymbol{\alpha}_i$ and $\mathbf{w}_{t(i)}$ is given by:

$$\begin{aligned} \ln P_i(\mathbf{r}_i|\boldsymbol{\alpha}_i) \;=\; & \sum_c r_{it(i)}^c \ln p_0\big(c\big|\, \boldsymbol{\alpha}_i, \mathbf{w}_{t(i)}\big) \\ & + \sum_{t=t(i)+1}^{T(i)} \sum_{c\in\mathcal{C}} \sum_{d\in\mathcal{C}} r_{it-1}^c r_{it}^d \ln p(d|c, \boldsymbol{\alpha}_i, \mathbf{w}_t). \end{aligned} \tag{9}$$

Equation (9) yields the contribution of the observations of farmer *i* to the likelihood function of the considered CMP choice model of our dataset, with $\boldsymbol{\alpha}_i = (\boldsymbol{\gamma}_i, \rho_i, \boldsymbol{\mu}_i, \sigma_i, \boldsymbol{\eta}_i)$. The endogeneity of CMP choices with respect to the crop production choices is assumed to only depend on unobserved farm heterogeneity that is explicitly modelled through random parameters $\boldsymbol{\alpha}_i$. This assumption implies that CMP choices on the one hand, and input use and yield levels on the other hand are independent conditional on $\boldsymbol{\alpha}_i$. This implies in turn that error terms $\mathbf{e}_{it} = (\mathbf{e}_{it}^{|c}|\, c \in \mathcal{C})$ and $\boldsymbol{\varepsilon}_{it}$ are independent and, finally, that the likelihood function related to $\mathbf{q}_{it}$ is easily obtained.

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

3.3. Overall identification strategy

Considering latent CMPs together with random parameters rises challenging identification issues. Our identification strategy mainly relies on three elements. (a) We design models that fully describe how CMP impact farmers' input use and yield levels and how farmers are supposed to choose the CMP they use. (b) Our models explicitly account for the effects of unobserved random events and unobserved heterogeneity in production conditions. (c) Finally, we exploit the information content of panel data that describe the input use and yield histories of a large sample of wheat producers. In particular, farms observed for less than four times in the panel were dropped for securing the identification of the model, which features a hidden Markovian process and farm-specific random parameters.

The effects of unobserved heterogeneity (e.g., soil quality) 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 CMP can evolve over time. Importantly, wheat prices display substantial variations during the observation period. They are at low levels from 1998 to 2006. They skyrocketed in 2007 and have remained relatively high, albeit quite volatile, since then. We expect these price variations to be sufficient for leading farmers to change their CMPs and, thus, for providing sufficient exogenous variations for disentangling the effects of CMP choices from those of production condition heterogeneity. The effects of unobserved random events are disentangled from those of CMP changes by explicitly modelling farmers' CMP choice sequences as Markov processes. While random events are assumed independent across years, CMP choices are expected to display smoother dynamic patterns due to factors preventing farmers to easily switch from a CMP to another.

Of course, the functional forms of the CMP choice probabilities and of the CMP specific yield and input use models, together with our distributional assumptions, also play a crucial role for guiding the estimation process and ensuring the model identification.

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 $\boldsymbol{\alpha}_i \sim \mathcal{N}(\boldsymbol{\alpha}_0, \boldsymbol{\Omega}_0)$, where variance matrix $\boldsymbol{\Omega}_0$ is left unrestricted. Since our model is fully parametric, we consider estimating its parameters, $\boldsymbol{\theta}_0 = (\mathbf{d}_0, \boldsymbol{\Delta}_0, \boldsymbol{\alpha}_0, \boldsymbol{\Omega}_0, \Sigma_0^c : c \in \mathcal{C})$, using the Maximum Likelihood (ML) approach. The contribution of farmer i to the sample likelihood function at $\boldsymbol{\theta}$ conditional on random parameter $\boldsymbol{\alpha}_i$ is given by:

$$\ell_i(\boldsymbol{\theta}|\boldsymbol{\alpha}_i) = \sum_{c_{it(i)} \in \mathcal{C}} \dots \sum_{c_{iT(i)} \in \mathcal{C}} p_0(c_{it}|\boldsymbol{\alpha}_i, w_{t(i)}) \prod_{t=t(i)+1}^{T(i)} p(c_{it}|c_{it-1}, \boldsymbol{\alpha}_i, w_t) \prod_{t=t(i)}^{T(i)} \varphi(\mathbf{u}_{it}^c; \boldsymbol{\Sigma}^{c_{it}}), \quad (10)$$

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:

$$\ell_i(\boldsymbol{\theta}) = \int \ell_i(\boldsymbol{\theta}|\boldsymbol{\varsigma})\, \varphi(\boldsymbol{\varsigma} - \boldsymbol{\alpha}\,; \boldsymbol{\Omega})\, d\boldsymbol{\varsigma}. \quad (11)$$

Computing likelihood terms $\ell_i(\boldsymbol{\theta})$ is particularly challenging but can be considered by relying on suitable extensions of the Expectation-Maximization (EM) algorithm of Dempster et al. (1977). 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.

Yet, as it stands in Equation (10) 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. Moreover, the integration problem involved in Equation (11) 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 $\boldsymbol{\theta}$. 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 (Kuhn and Lavielle, 2005; Lavielle and Mbogning, 2014). This algorithm relies on a stochastic approximation approach for solving the E step. Detailed description of the E step, its SAEM version, and the M step are provided in Appendix 9.1.

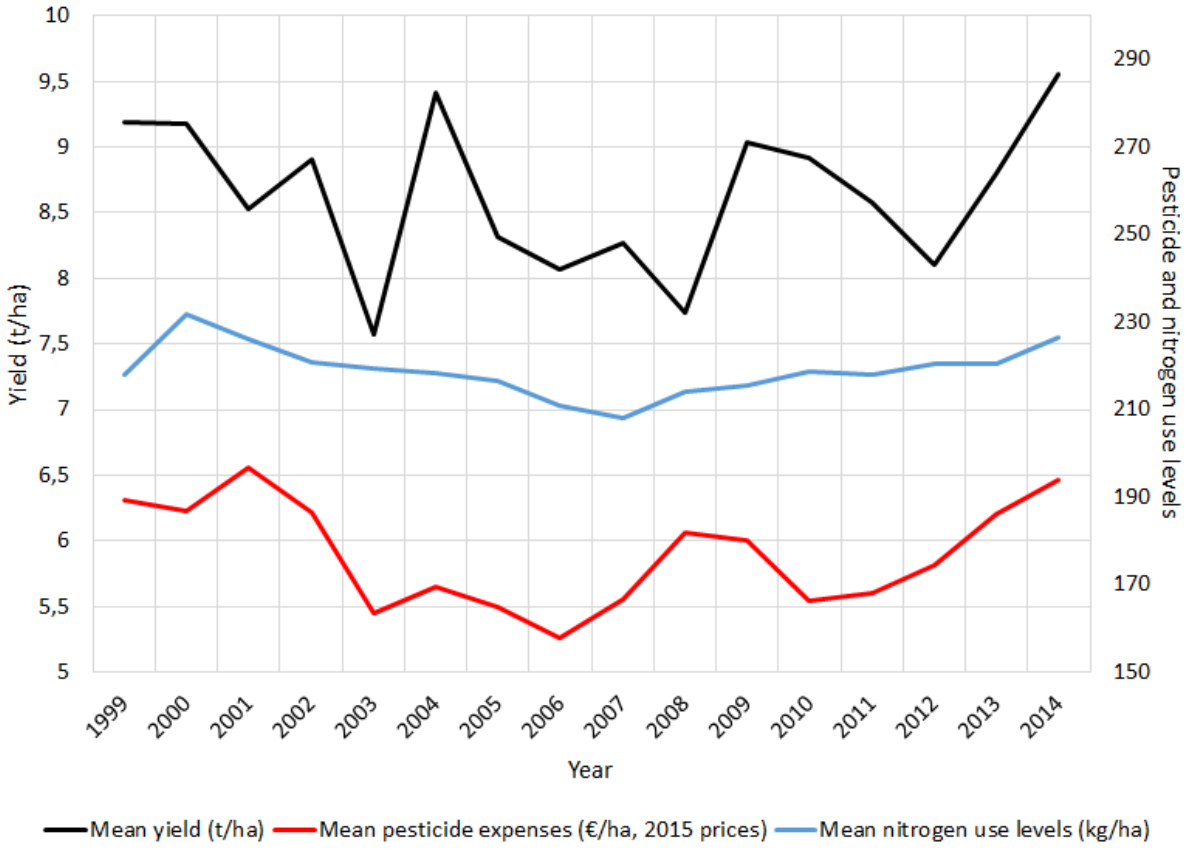
5. Data

We use an unbalanced panel dataset that considers, from 1998 to 2014, input use and yield levels of winter wheat for a sample of farmers located in the *La Marne département*, a French territorial division similar to a small US county. These data were extracted from cost accounting data provided by an accountancy agency. Overall, the dataset gathers 1351 farmers that are observed for 10 years on average.

Our application primarily makes use of recorded winter wheat yields, fertilizer uses and

pesticide expenditures devoted to wheat production. Mean yield, nitrogen fertilizer uses and pesticide expenses are reported in Figure 1. Pesticide uses depend on pest and weed pressures, which themselves depend on weather conditions, implying that pesticide uses display greater variability than fertilizer uses. Still, wheat yields display the highest variability across time, due to their strong responsiveness to weather conditions. They vary from 7.57t/ha in 2003 to 9.55t/ha in 2014 with an overall mean around 8.6t/ha.

Figure 1. Annual mean yields and uses of agrochemicals, north eastern France, 1999–2014

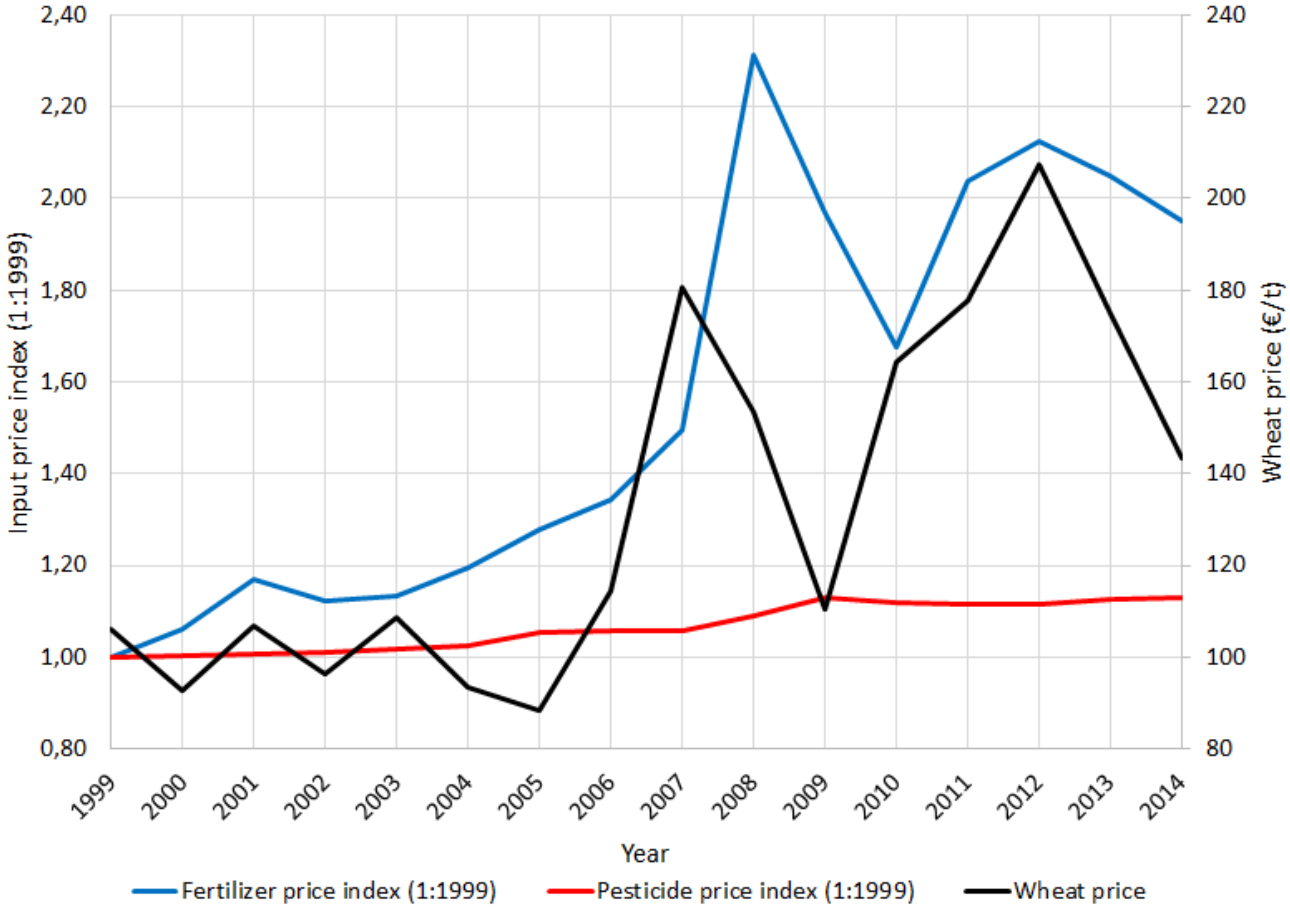


Source: Authors' calculations.

Assuming that farmers' wheat price anticipations are naïve (Koutchadé et al., 2018, 2021), we used the lagged wheat prices observed at the farm gate as the wheat prices expected by farmers. Pesticide price indices were obtained from the French department of Agriculture (Agreste, INSEE, 2015). As for fertilizer prices, we rely on the cost accounting data to build a

proxy for nitrogen prices.[10] Figure 2 presents the considered price levels. Pesticide prices are quite steady whereas wheat and nitrogen prices display similar upward trends.

Figure 2. Wheat, pesticide and nitrogen annual prices, France, 1999–2014



Source: IPAMPA data from Agreste/INSEE for pesticides prices; authors' calculations (wheat and mineral nitrogen prices).

Wheat and nitrogen prices significantly increased after the 2007 boom and have been relatively volatile since then. Given the economic context prevailing from 1999 to 2007, *i.e.* low wheat prices, and the promotion process of low-input CMPs during this period, we expect a small share of farmers using low-input in the late 1990s and an increase in this share

[10] This proxy is obtained by regressing the detailed fertilizer uses (*i.e.*, uses of N, P and K) on global fertilizer expenses. The resulting coefficients are used as proxies for nitrogen, phosphorus and potassium prices. Our application only considers nitrogen uses. Nitrogen fertilizers, which represent more than 50% of fertilizer uses, are used as short run inputs while phosphorus and potassium are used in a long run perspective.

until 2006.

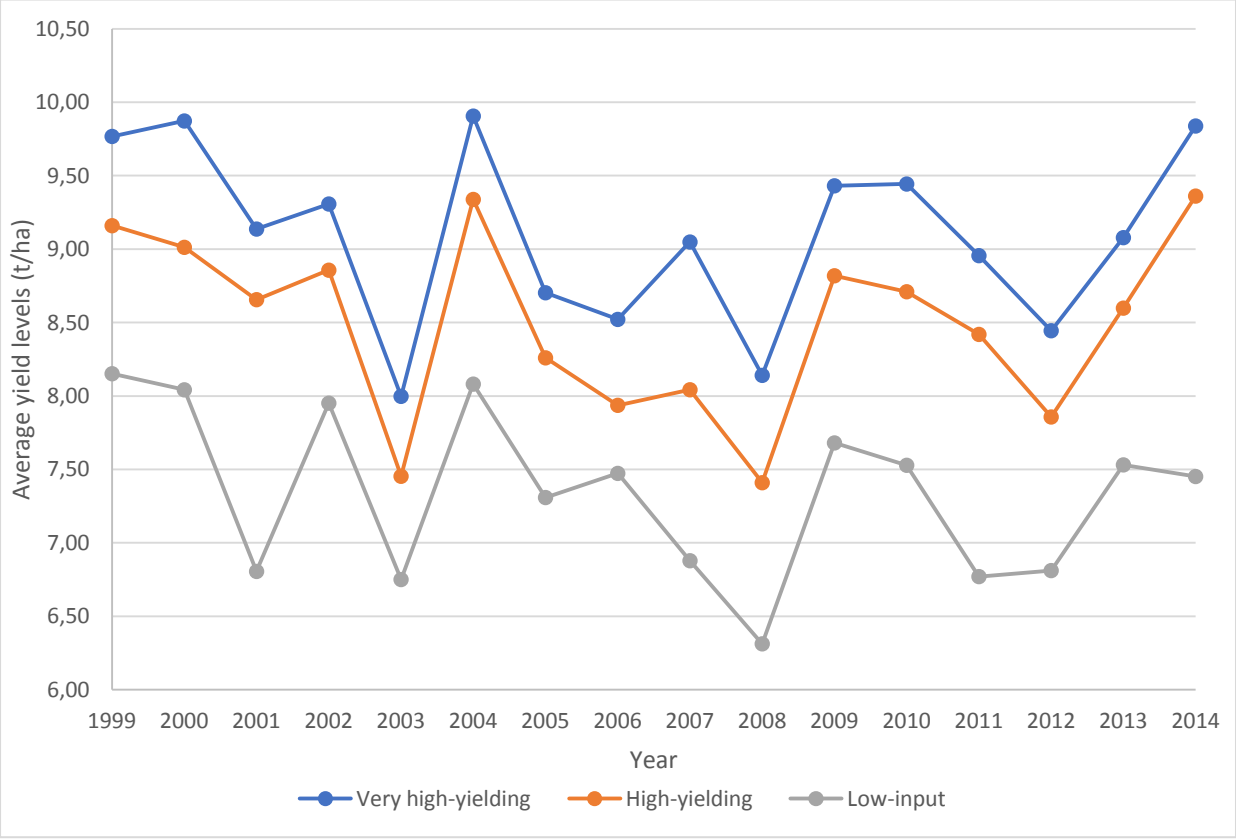
6. Results

In this section we present the results from the RPHM model. Estimates of the parameters of the probability distribution of the random parameters of the CMP choice probability functions, the production functions and the inputs demand model can be found in Appendix B. Overall, the estimated standard deviations of the ML estimator the RPHM model shows that its parameters are precisely estimated with the considered dataset. This observation holds for both the fixed parameters and the – mean, variance and covariance – parameters of the probability distribution of its random parameter vector. In particular, the random parameter variance estimates appear to be significantly non-null, thereby confirming the presence of unobserved heterogeneity effects in the considered RPHM model.

The estimated RPHM model distinguishes three CMP types. Very-high-yielding CMPs are associated to the largest levels of yield and input use. High-yielding CMPs are characterized by yield and input use levels slightly lower than their counterparts in very-high-yielding CMPs. Low-input CMPs display the lowest yield and input use levels. The mean yield series from 1999 to 2014 – respectively the mean input uses series – observed in each of those three CMP types are depicted in Figure 3 – respectively Figures 4 and 5. On average, high-yielding CMPs use 1.8% less nitrogen and 16.5% less pesticide than very high-yielding CMPs for a 6.7% decrease in yields. Low-input CMPs use on average 3.4% (respectively 5.2%) less

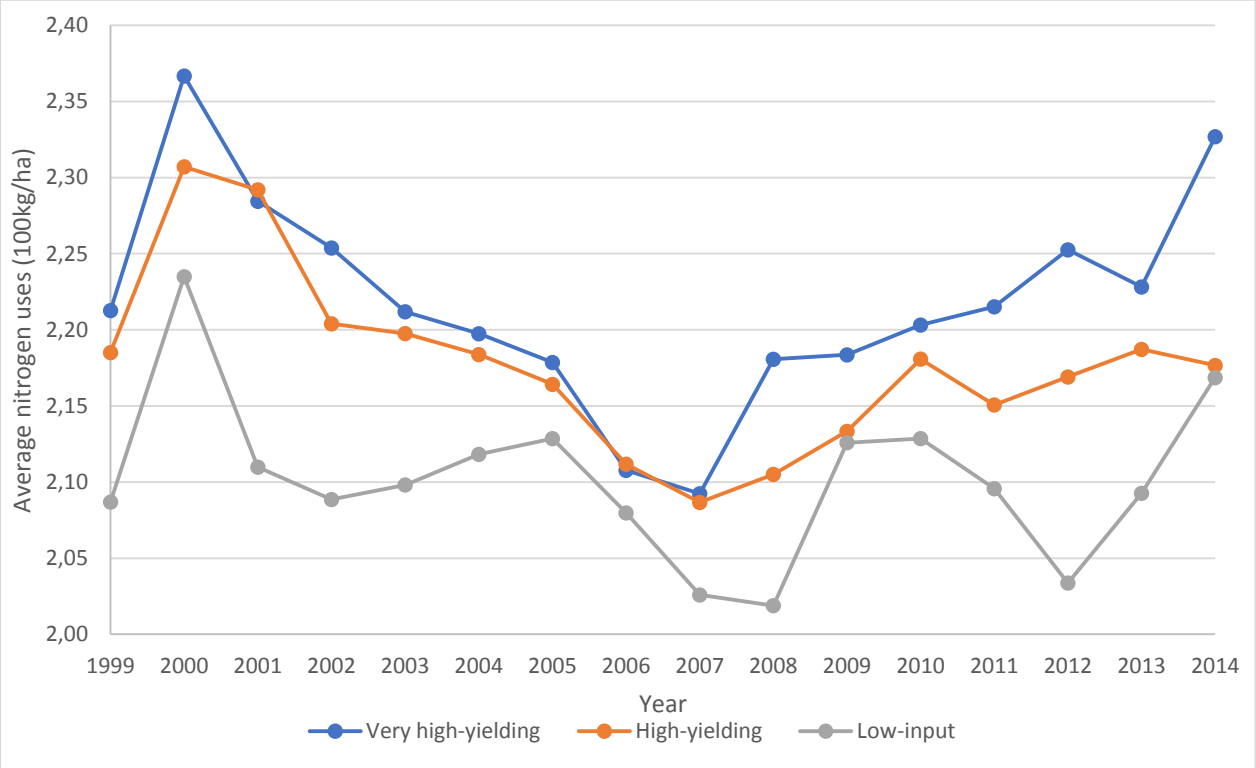
nitrogen, 15.4% (respectively 29.3%) less pesticides than high-yielding CMPs (respectively very high-yielding CMPs) for an average decrease of 13.4% (respectively 19.2%) in yields. These results tend to show that our RPHM model uncovers CMP types close to the ‘‘maximum yield’’, ‘‘conventional’’ and ‘‘low-input/multi-resistant varieties’’ CMPs considered in the experiments presented by Rolland et al. (2003) and Loyce et al. (2012).

Figure 3. Annual mean yield per CMP type, north eastern France, 1999–2014



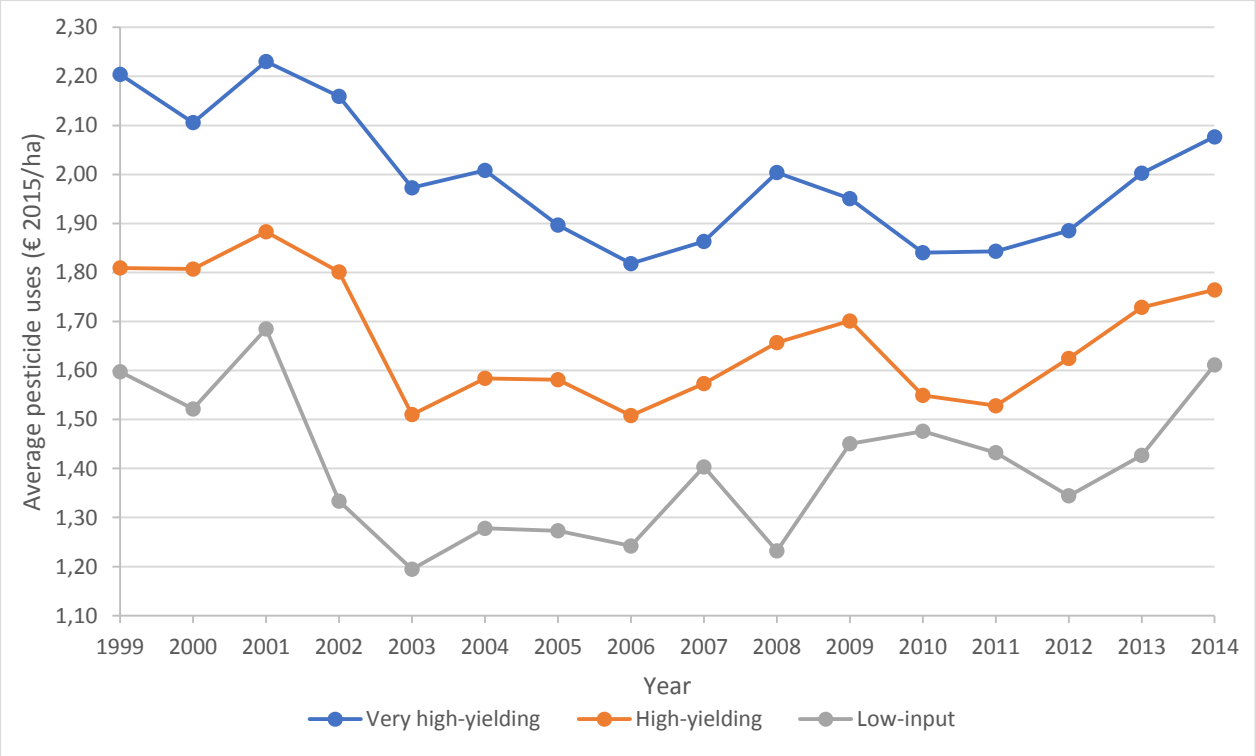
Source: Authors’ calculations.

Figure 4. Annual mean nitrogen use per CMP type, north eastern France, 1999–2014



Source: Authors' calculations.

Figure 5. Annual mean pesticide use per CMP type, north eastern France, 1999–2014



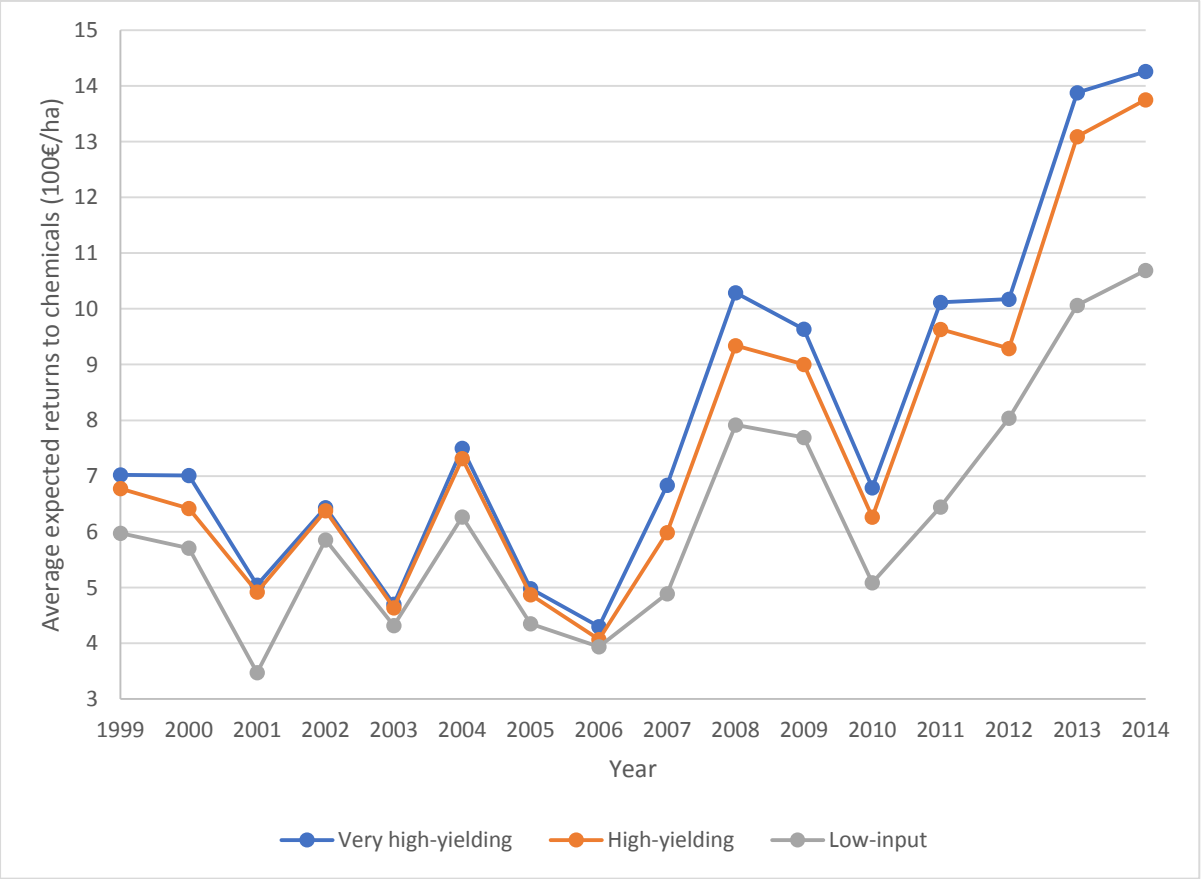
Source: Authors' calculations.

Another result supports this conclusion. Fungicide is the pesticide category the use of which varies the most across CMP types.[11] This was expected as disease pressure strongly increases with production intensity (Meynard, 1985, 1991). In particular, low-input CMPs, which were proposed by Meynard (1985, 1991) and tested by Rolland et al. (2003) and Loyce et al. (2008, 2012) mostly target reductions in fungicide and, to a lesser extent, in fertilizer uses. Farmers using low-input CMPs use wheat varieties resistant to diseases, lower their sowing density and slightly postpone their sowing date. These techniques lower disease pressure and yield levels, thereby lowering fungicide and nitrogen use requirements. The impacts of these techniques on insect infestations and weed pressure are limited.

Differences in input uses and in yields impact the expected returns (to chemicals) associated to each CMP category. Figure 6 shows that year 2007 is a turning point for expected return levels. Wheat prices suddenly increased in 2007 and remained relatively high, on average, since then. Differences in the expected returns of the considered CMP categories are limited when expected wheat price are relatively low. Yet, returns of low-input CMPs tend to be slightly smaller than those of high-yielding CMPs at these price levels. High wheat price levels tend to magnify the effects of yield differences on CMP returns, implying that high wheat prices tend to make (very) high-yielding CMPs (much) more profitable than low-input CMPs.

[11] Detailed results for fungicides, insecticides and herbicides can be found in Supplemental Appendix 10.1.

Figure 6: Estimated annual mean expected return per CMP type, north eastern France, 1999–2014

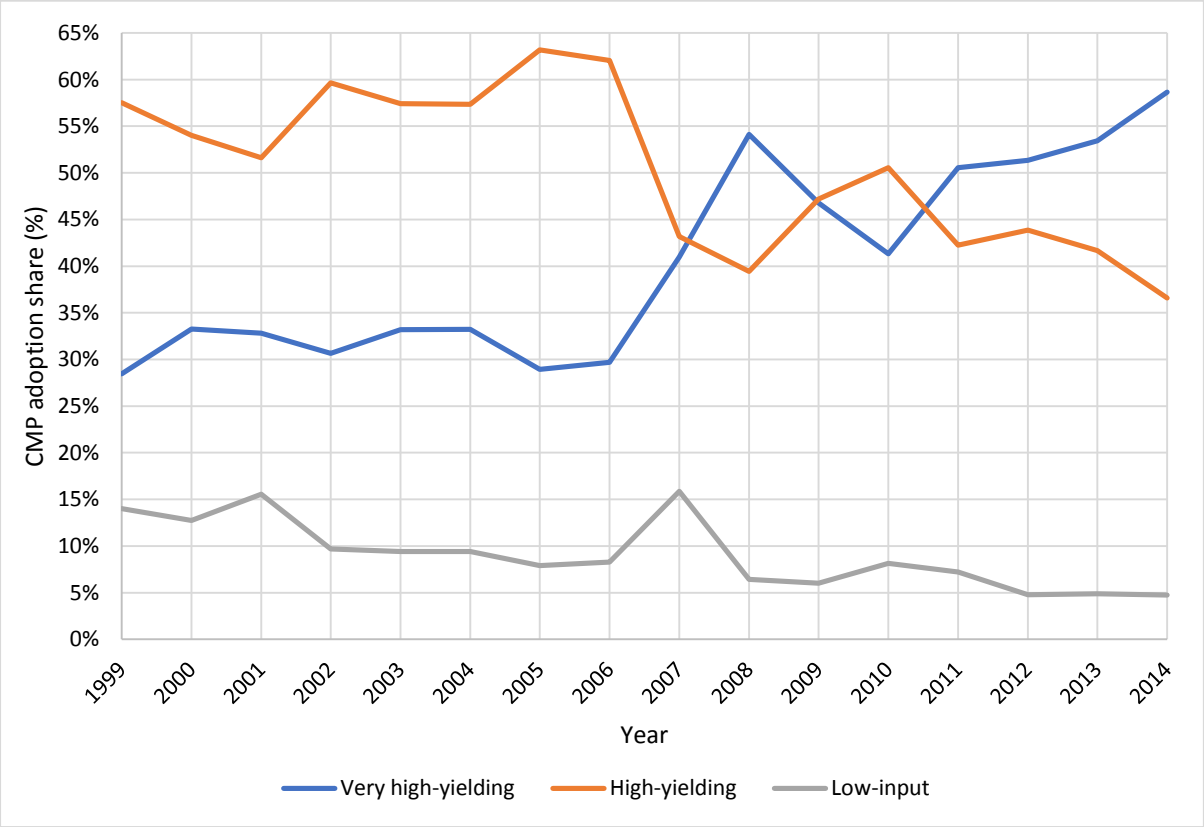


Source: Authors' calculations.

The estimated shares of farmers of our sample using the considered CMP categories is depicted in Figure 7. Until the 2007 crisis these shares are relatively steady with approximately 10% of farmers using low-input CMPs, 60% using high-yielding techniques and around 30% using very high-yielding techniques. The estimated share of farmers using low-input CMPs declined by steps during the considered period. Around 15% of farmers are estimated to use low-input CMPs from 1999 to 2001. The low-input CMP share then dropped to 10% in 2002 and remained at this level until 2006. It dropped once again after 2007 at 7% in 2008 and progressively declined to around 5% in 2014. The most drastic changes concern the shares of high-input and high-yielding CMPs with a sharp increase in the use of very high-yielding techniques in 2007/2008. The high wheat prices observed after 2007 tend to

favor production practices aimed to achieve yield levels that are close to local yield potentials.

Figure 7: Estimated annual share of farmers per CMP type, north eastern France, 1999–2014



Source: Authors' calculations.

Low-input practices being more profitable when crop prices are low is pointed out in the agronomic literature (Loyce et al., 2012; Rolland et al., 2003). Yet, in our case, the disadoption of low-input CMPs due to the drastic increase in wheat price that occurred after 2007 is relatively limited, suggesting that adoption of these CMPs obey to economic reasons but is also linked to farmer personal values as environmental and societal concerns (Frey and Stutzer, 2006; Howley et al., 2015; Mzoughi, 2011). This result is confirmed by the analysis of the estimated parameters of the CMP choice model. Indeed, according to our estimates, switching to the low-input technology entails relatively high transition costs for farmers using

high-yielding or very high-yielding CMPs. Also, the initial cost for adopting the low-input CMP is higher than for the very-high-yielding and high-yielding CMPs. Yet, farmers using a low-input CMP tend to stick to this technology, as is the case for the other CMP types. These results may also indicate that low-input CMPs did not have sufficient time to diffuse in the farmers' population. Indeed, the time window that favored their adoption was very short. It only spanned the very late 90's and, more surely, the early 00's.

The lack of economic incentives for adopting low-input CMPs is confirmed by simulation results obtained from the estimated model.[12] We find a very limited responsiveness of CMP choices to pesticide taxes. Even a 100% tax is insufficient to fill the gap between the low-input CMP returns and their high-yielding counterparts. The chemical input savings permitted by low-input CMPs are too limited – and/or that low-input CMPs entail yield reductions that are too large – for input taxes to create sufficient incentives towards these environmentally friendly CMPs. This limited responsiveness of input uses to prices is in line with the literature showing low price elasticity of pesticides. Yet, while previous studies considered short run elasticities – holding production technology fixed – our study tends to demonstrate that medium run elasticities – allowing changes in CMPs – of pesticide uses are also small (in absolute value).

7. Conclusion

We proposed a random parameter model with endogenous regimes that follow a hidden Markov chain for uncovering latent CMPs and their associated yield and input use levels. This model explicitly considers that (latent) CMP expected returns (to chemical inputs)

[12] Detailed results from this simulation work can be found in Appendix 9.3.

impact farmers' CMP choices. Its random parameters account for farmers and farm unobserved heterogeneity. Farmers' CMP choice is defined as a Markov process for accounting for eventual CMP switching costs that may underlie farmers' often observed reluctance to change their production practices.

Our application on a panel dataset describing the wheat production choices of farmers allowed us to uncover three CMP types used in the Marne area from 1999 to 2014. Very high-yielding and high-yielding CMPs are standard, or conventional, crop production practices in this area. By contrast, low-input CMPs are innovative production practices designed by agronomists for reducing chemical input uses, especially pesticide uses, by lowering the target yield levels (Loyce et al., 2008, 2012; Meynard, 1991). Importantly, the characteristics of the uncovered low-input CMPs are very close to those tested by agronomists in the Marne area during the mid-1990s (Larédo and Hocdé, 2014). 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 enables 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 high-yielding CMPs to very 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, but they did not fully deter the use of low-input CMPs. This suggests that non-economic motives impact farmers' production choices, at least those adopting low-input practices. Non-financial drivers of

farmers' choices may include attitude toward the environment, health concerns and taste for agronomy, which were shown play a significant role in the adoption of environmentally friendly agricultural production practices (Howley et al., 2015). Yet, we are not able to disentangle the effects of this wide variety of CMP choice drivers and analyze their effects in farmers' production choices. Thus, our ability to provide insights on public policies, aimed to foster the adoption of chemical input saving production practices by farmers, is limited to purely economic drivers.

Lastly, our simulation work tends to show that chemical input uses differences between low-input and more conventional CMPs are too small for taxes on chemical inputs to imply large relative profitability effects. Indeed, high rate taxes on chemical inputs may rather 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. More generally, our results tend to highlight the economic efficiency of combining high-yielding CMPs and chemical crop protection and, as a result, provide support to the viewpoint stating that current agricultural production strongly depends on chemical crop protection given the current economic context and currently available production technologies. Indeed, conventional CMPs are designed while taking for granted that sufficiently high levels of pesticide uses enable farmers to achieve high crop protection levels at relatively low cost. Pesticide saving CMPs are conceived by implicitly setting constraining upper bounds on the uses of inputs that are both technically and economically efficient and by accepting reductions in target yield levels (for containing pest and disease pressure). These low-input CMPs are put at a disadvantage as long as pesticides are available at relatively low cost.

If our Markovian modelling framework allows us to account for the dynamic aspect of CMP

choice, the specific features of the diffusion dynamics of agricultural production technologies are ignored. Learning-by-doing and learning-by-others mechanisms are empirically documented by economists, especially in developing countries (Chavas and Nauges, 2020; Foster and Rosenzweig, 2010; Marra et al., 2003). These features are largely overlooked in our modelling framework. The fact that the adoption of the considered technologies is unobserved in our case makes it particularly difficult to account for them. Yet we expect that the effects of the congruent learning processes to be limited given the low adoption rate of low-input CMPs in France.

Of course, our modelling framework could be improved. Two limitations appear to be particularly worth addressing. First, the initial CMP adoption probabilities of our Markov model 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. Second, further structuring the latent yield and input use models considered in our modelling framework could also be fruitful. This would enable the investigation of drivers of input uses conditionally on predetermined CMP choices.

Applying the approach proposed in this article to other crops is also of interest. Low-input CMPs have not been explicitly designed and promoted by agronomists for crops other than wheat. Yet, farmers using low-input CMPs for wheat production may extend 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, 2021). 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.

8. References

Agreste, INSEE. (2015). *Indices des prix d'achat des moyens de production agricole (IPAMPA)* [Data set]. https://agreste.agriculture.gouv.fr/agreste-saiku/?plugin=true&query=query/open/D_0049#query/open/D_0049 , last accessed 14 June 2023.

Aubertot, J.-N., Barbier, J. M., Carpentier, A., Gril, J.-N., Guichard, L., Lucas, P., Savary, S., and Voltz, M. (2007). *Pesticides, agriculture et environnement. Réduire l'utilisation des pesticides et en limiter les impacts environnementaux. Expertise scientifique collective Inra-Cemagref (décembre 2005)*. Quae , 120 p.

Böcker, T. G., and Finger, R. (2017). A Meta-Analysis on the Elasticity of Demand for Pesticides. *Journal of Agricultural Economics 68*: 518–533.

Booth, J. G., and Hobert, J. P. (1999). Maximizing generalized linear mixed model likelihoods with an automated Monte Carlo EM algorithm. *Journal of the Royal Statistical Society: Series B (Statistical Methodology) 61*: 265–285.

Booth, J. G., Hobert, J. P., and Jank, W. (2001). A survey of Monte Carlo algorithms for maximizing the likelihood of a two-stage hierarchical model. *Statistical Modelling 1*: 333–349.

Chavas, J.-P., and Holt, M. T. (1990). Acreage Decisions Under Risk: The Case of Corn and Soybeans. *American Journal of Agricultural Economics 72*: 529–538.

Chavas, J.-P., and Nauges, C. (2020). Uncertainty, Learning, and Technology Adoption in Agriculture. *Applied Economic Perspectives and Policy 42*: 42–53.

Dakpo, K. H., Latruffe, L., Desjeux, Y., and Jeanneaux, P. (2021). Latent Class Modelling for a Robust Assessment of Productivity: Application to French Grazing Livestock Farms. *Journal of Agricultural Economics 72*: 760–781.

Delattre, M., and Lavielle, M. (2012). Maximum likelihood estimation in discrete mixed hidden Markov models using the SAEM algorithm. *Computational Statistics & Data Analysis 56*: 2073–2085.

Delyon, B., Lavielle, M., and Moulines, E. (1999). Convergence of a Stochastic Approximation Version of the EM Algorithm. *The Annals of Statistics 27*: 94–128.

Dempster, A. P., Laird, N. M., and Rubin, D. B. (1977). Maximum Likelihood from Incomplete Data Via the EM Algorithm. *Journal of the Royal Statistical Society: Series B (Methodological) 39*: 1–22.

Femenia, F., and Letort, E. (2016). How to significantly reduce pesticide use: An empirical evaluation of the impacts of pesticide taxation associated with a change in cropping practice. *Ecological Economics 125*: 27–37.

Foster, A. D., and Rosenzweig, M. R. (2010). Microeconomics of Technology Adoption. *Annual Review of Economics 2*: 10.1146/annurev.economics.102308.124433.

Frey, B. S., and Stutzer, A. (2006). *Environmental Morale and Motivation* , SSRN Scholarly Paper No. 900370. Rochester, NY.:

Frisvold, G. B. (2019). How low can you go? Estimating impacts of reduced pesticide use. *Pest Management Science 75*: 1223–1233.

Genius, M., Koundouri, P., Nauges, C., and Tzouvelekas, V. (2014). Information Transmission in Irrigation Technology Adoption and Diffusion: Social Learning, Extension Services, and Spatial Effects. *American Journal of Agricultural Economics 96*: 328–344.

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

Griliches, Z. (1957). Hybrid Corn: An Exploration in the Economics of Technological Change. *Econometrica 25*: 501–522.

Gu, M. G., and Zhu, H.-T. (2001). Maximum Likelihood Estimation for Spatial Models by Markov Chain Monte Carlo Stochastic Approximation. *Journal of the Royal Statistical Society Series B: Statistical Methodology 63*: 339–355.

Howley, P., Buckley, C., O Donoghue, C., and Ryan, M. (2015). Explaining the economic 'irrationality' of farmers' land use behaviour: The role of productivist attitudes and non-pecuniary benefits. *Ecological Economics 109*: 186–193.

Koutchadé, O. P., Carpentier, A., and Femenia, F. (2018). Modeling Heterogeneous Farm Responses to European Union Biofuel Support with a Random Parameter Multicrop Model. *American Journal of Agricultural Economics 100*: 434–455.

Koutchadé, O. P., Carpentier, A., and Femenia, F. (2021). Modeling Corners, Kinks, and Jumps in Crop Acreage Choices: Impacts of the EU Support to Protein Crops. *American Journal of Agricultural Economics 103*: 1502–1524.

Kuhn, E., and Lavielle, M. (2005). Maximum likelihood estimation in nonlinear mixed effects models. *Computational Statistics & Data Analysis 49*: 1020–1038.

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

Lavielle, M., and Mbogning, C. (2014). An improved SAEM algorithm for maximum likelihood estimation in mixtures of non linear mixed effects models. *Statistics and Computing 24*: 693–707.

Los, E., Gardebroek, C., and Huirne, R. (2021). Firm-specific responses to energy policies in Dutch horticulture. *European Review of Agricultural Economics 48*: 362–384.

Loyce, C., Meynard, J. M., Bouchard, C., Rolland, B., Lonnet, P., Bataillon, P., Bernicot, M. H., Bonnefoy, M., Charrier, X., Debote, B., Demarquet, T., Duperrier, B., Félix, I., Heddadj, D., Leblanc, O., Leleu, M., Mangin, P., Méausoone, M., and Doussinault, G. (2008). Interaction between cultivar and crop management effects on winter wheat diseases, lodging, and yield. *Crop Protection 27*: 1131–1142.

Loyce, C., Meynard, J. M., Bouchard, C., Rolland, B., Lonnet, P., Bataillon, P., Bernicot, M. H., Bonnefoy, M., Charrier, X., Debote, B., Demarquet, T., Duperrier, B., Félix, I., Heddadj, D., Leblanc, O., Leleu, M., Mangin, P., Méausoone, M., and Doussinault, G. (2012). Growing winter wheat cultivars under different management intensities in France: A multicriteria assessment based on economic, energetic and environmental indicators. *Field Crops Research 125*: 167–178.

Mahé, L. P., and Rainelli, P. (1987). Impact des pratiques et des politiques agricoles sur l'environnement. *Cahiers d'économie et sociologie rurales 4*: 9.

Marra, M., Pannell, D. J., and Abadi Ghadim, A. (2003). The economics of risk, uncertainty and learning in the adoption of new agricultural technologies: where are we on the learning curve? *Agricultural Systems 75*: 215–234.

Martinez Cillero, M., Thorne, F., Wallace, M., and Breen, J. (2019). Technology heterogeneity and policy change in farm-level efficiency analysis: an application to the Irish beef sector. *European Review of Agricultural Economics 46*: 193–214.

Maruotti, A. (2011). Mixed Hidden Markov Models for Longitudinal Data: An Overview. *International Statistical Review 79*: 427–454.

Meynard, J. M. (1985). *Construction d'itinéraires techniques pour la conduite du blé d'hiver* [Phdthesis, Institut National Agronomique Paris Grignon]. https://hal.inrae.fr/tel-02856778 , last accessed 11 January 2023.

Meynard, J. M. (1991). *Pesticides et itineraires techniques*.

Meynard, J. M., Rolland, B., Loyce, C., Félix, I., and Lonnet, P. (2009). Quelles combinaisons variétés/conduites pour améliorer les performances économiques et environnementales de la culture de blé tendre ? *Innovations Agronomiques 7*: 29.

Michler, J. D., Tjernström, E., Verkaart, S., and Mausch, K. (2019). Money Matters: The Role of Yields and Profits in Agricultural Technology Adoption. *American Journal of Agricultural Economics 101*: 710–731.

Monolix methodology. (2014). Lixoft. http://lixoft.com/wp-content/uploads/2016/03/monolixMethodology.pdf , last accessed 3 April 2023.

Mzoughi, N. (2011). Farmers adoption of integrated crop protection and organic farming: Do moral and social concerns matter? *Ecological Economics 70*: 1536–1545.

Oerke, E.-C. (2006). Crop losses to pests. *The Journal of Agricultural Science 144*: 31–43.

Orea, L., and Kumbhakar, S. C. (2004). Efficiency measurement using a latent class stochastic frontier model. *Empirical Economics 29*: 169–183.

Popp, J. (2011). Cost-benefit analysis of crop protection measures. *Journal Für Verbraucherschutz Und Lebensmittelsicherheit 6*: 105–112.

Rabiner, L. R. (1989). A tutorial on hidden Markov models and selected applications in speech recognition. *Proceedings of the IEEE 77*: 257–286.

Renner, S., Sauer, J., and El Benni, N. (2021). Why considering technological heterogeneity is important for evaluating farm performance? *European Review of Agricultural Economics 48*: 415–445.

Rolland, B., Bouchard, C., Loyce, C., Meynard, J. M., Guyomard, H., Lonnet, P., and Doussinault, G. (2003). Technical systems for wheat using low input quantities and rustic varieties : a solution to conciliate economy with environment
[Des itinéraires techniques à bas niveaux d'intrants pour des variétés rustiques de blé

tendre : une alternative pour concilier économie et environnement]. *Post-Print* Article hal-02681424.

Ruud, P. A. (1991). Extensions of estimation methods using the EM algorithm. *Journal of Econometrics 49*: 305–341.

Skevas, T., Oude Lansink, A. G. J. M., and Stefanou, S. E. (2013). Designing the emerging EU pesticide policy: A literature review. *NJAS - Wageningen Journal of Life Sciences 64–65*: 95–103.

Suri, T. (2011). Selection and Comparative Advantage in Technology Adoption. *Econometrica 79*: 159–209.

Wei, G. C. G., and Tanner, M. A. (1990). A Monte Carlo Implementation of the EM Algorithm and the Poor Man's Data Augmentation Algorithms. *Journal of the American Statistical Association 85*: 699–704.

9. Appendices

9.1. The estimation procedure

Let define vectors $\boldsymbol{q}_{(i)} = (\boldsymbol{q}_{it} : t \in H_i)$ and $\boldsymbol{z}_{(i)} = (\boldsymbol{z}_{it} : t \in H_i)$. The complete data of our model consists of (*i*) the observed variable vectors $\boldsymbol{\zeta}_i = (\boldsymbol{q}_{(i)}, \boldsymbol{w}_{(i)}, \boldsymbol{z}_{(i)})$, (*ii*) the latent CMP choices sequence $\boldsymbol{r}_{(i)}$ and (*iii*) the random parameters vector $\boldsymbol{\alpha}_i$, for $i = 1, \ldots, N$. The complete data log-likelihood function is the sample log-likelihood function of the joint model of the dependent and missing variables, $(\boldsymbol{q}_{(i)}, \boldsymbol{r}_{(i)}, \boldsymbol{\alpha}_i)$, given the exogenous variables of the model, $(\boldsymbol{w}_{(i)}, \boldsymbol{z}_{(i)})$. The complete data log-likelihood function at $\boldsymbol{\theta}$ of our model is given by:

$$\ln \mathrm{L}^{\mathrm{C}}(\boldsymbol{\theta}) = \sum_{i=1}^{N} \ln \ell_i^{C}(\boldsymbol{\theta} \mid \mathbf{r}_{(i)}, \boldsymbol{\alpha}_i), \tag{12}$$

where

$$\ln \ell_i^{C}(\boldsymbol{\theta} \mid \mathbf{r}_{(i)}, \boldsymbol{\alpha}_i) =$$

$$\begin{Bmatrix} \sum_c r_{it(i)}^{c} \ln p_0(c \mid \boldsymbol{\alpha}_i, \boldsymbol{w}_{t(i)}) + \sum_{t=t(i)+1}^{T(i)} \sum_{c \in \mathcal{C}} \sum_{d \in \mathcal{C}} r_{it-1}^{c} r_{it}^{d} \ln p(d|c, \boldsymbol{\alpha}_i, \boldsymbol{w}_t) \\ + \sum_{t=t(i)}^{T(i)} \sum_{c \in \mathcal{C}} r_{it}^{c} \ln \varphi(\boldsymbol{q}_{it} - \boldsymbol{b}_i^{c} - \boldsymbol{d}_t - \boldsymbol{\Delta} \boldsymbol{z}_{it};\, \boldsymbol{\Sigma}^{c}) + \ln \varphi(\boldsymbol{\alpha}_i - \boldsymbol{\alpha}\,;\, \boldsymbol{\Omega}) \end{Bmatrix}.$$

At iteration n of an EM type algorithm, the objective of the E step is to integrate $\ln \mathrm{L}^{\mathrm{C}}(\boldsymbol{\theta})$ over the probability distribution of the missing data $(\boldsymbol{r}_{(i)}, \boldsymbol{\alpha}_i)$, conditional on the observed data $\boldsymbol{\zeta}_i$ evaluated at $\boldsymbol{\theta}^{(\mathrm{n})}$, the last available estimate of $\boldsymbol{\theta}_0$. This consists of computing the conditional expectations $E^{(n)}\left[\ln \ell_i^{C}(\boldsymbol{\theta} \mid \boldsymbol{r}_{(i)}, \boldsymbol{\alpha}_i) | \boldsymbol{\zeta}_i\right]$, for $i = 1, \ldots, N$, where $E^{(n)}$ corresponds to the expectation calculated at iteration n given the current estimates $\boldsymbol{\theta}^{(\mathrm{n})}$.

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

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

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

$$E^{(n)}\left[\ln \ell_i^C(\boldsymbol{\theta} \mid \boldsymbol{r}_{(i)}, \boldsymbol{\alpha}_i)|\boldsymbol{\zeta}_i\right] =$$

$$\left\{\begin{array}{c} E^{(n)}\left[R\left(\boldsymbol{\alpha}_i, \boldsymbol{w}_{(i)}\right)|\boldsymbol{\zeta}_i\right] \\ + \sum_{t=t(i)}^{T(i)} \sum_{c \in \mathcal{C}} E^{(n)}\left[s_t\left(d, c|\boldsymbol{\alpha}_i, \boldsymbol{w}_{(i)}\right) \ln \varphi(\boldsymbol{q}_{it} - \boldsymbol{b}_i^c - \boldsymbol{d}_t - \boldsymbol{\Delta z}_{it};\ \boldsymbol{\Sigma}^c)|\ \boldsymbol{\zeta}_i\right] \\ + E^{(n)}[\ln \varphi(\boldsymbol{\alpha}_i - \boldsymbol{\alpha}\ ;\boldsymbol{\Omega}\ )\ |\boldsymbol{\zeta}_i] \end{array}\right\}, \quad (13)$$

where term: $R\left(\boldsymbol{\alpha}_i, \boldsymbol{w}_{(i)}\right) = \left\{\begin{array}{c} \sum_{c \in \mathcal{C}} p_0\left(c|\boldsymbol{\alpha}_i, \boldsymbol{w}_{t(i)}\right) \ln p_0\left(c|\boldsymbol{\alpha}_i, \boldsymbol{w}_{t(i)}\right) \\ + \sum_{t=t(i)}^{T(i)} \sum_{c \in \mathcal{C}} \sum_{d \in \mathcal{C}} s_t\left(d, c|\boldsymbol{\alpha}_i, \boldsymbol{w}_{(i)}\right) \ln p(d|c, \boldsymbol{\alpha}_i, \boldsymbol{w}_t) \end{array}\right\}$, does not involve any element of $\boldsymbol{\theta}$.

The expectations conditional on the observed data $\boldsymbol{\zeta}_i$ involved in Equation (13) 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 $\widehat{\boldsymbol{\alpha}}_i^{(n),j}$ their related weights, for $j = 1, \ldots, J^{(n)}$, where $J^{(n)}$ denote the draw number considered for iteration n.

The conditional expectation of $\ln \mathrm{L}^{\mathrm{C}}(\boldsymbol{\theta})$ is approximated by:

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

$$\sum_{i=1}^{N}\left\{\begin{array}{c} E^{(n)}\left[R\left(\boldsymbol{\alpha}_i, \boldsymbol{w}_{(i)}\right)|\boldsymbol{\zeta}_i\right] \\ +\sum_{t=t(i)}^{T(i)}\sum_{c\in\mathcal{C}}\sum_{j=1}^{J^{(n)}} \widehat{\omega}_i^{(n),j} s_t\left(d, c|\widehat{\boldsymbol{\alpha}}_i^{(n),j}, \boldsymbol{w}_{(i)}\right)\ln\varphi\left(\boldsymbol{q}_{it}-\widehat{\boldsymbol{b}}_i^{c,(n),j}-\boldsymbol{d}_t-\boldsymbol{\Delta z}_{it};\, \boldsymbol{\Sigma}^c\right) \\ +\sum_{j=1}^{J^{(n)}} \widehat{\omega}_i^{(n),j}\ln\varphi\left(\widehat{\boldsymbol{\alpha}}_i^{(n),j}-\boldsymbol{\alpha}\,;\boldsymbol{\Omega}\right) \end{array}\right\} \quad (14)$$

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

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

$$\max_{(\boldsymbol{d},\boldsymbol{\Delta},\boldsymbol{\Sigma})} \sum_{i=1}^{N}\sum_{t=t(i)}^{T(i)}\sum_{c\in\mathcal{C}}\sum_{j=1}^{J^{(n)}} \widehat{\omega}_i^{(n),j} s_t\left(d, c|\widehat{\boldsymbol{\alpha}}_i^{(n),j}, \boldsymbol{w}_{(i)}\right)\ln\varphi\left(\boldsymbol{q}_{it}-\widehat{\boldsymbol{b}}_i^{c,(n),j}-\boldsymbol{d}_t - \boldsymbol{\Delta z}_{it};\, \boldsymbol{\Sigma}^c\right), \quad (15)$$

yields $\left(\boldsymbol{d}^{(n+1)}, \boldsymbol{\Delta}^{(\mathrm{n}+1)}, \boldsymbol{\Sigma}^{\mathrm{c},(\mathrm{n}+1)} : c \in \mathcal{C}\right)$, the first part of $\boldsymbol{\theta}^{(\mathrm{n}+1)}$. On the other hand, solving problem:

$$\max_{(\boldsymbol{\alpha},\boldsymbol{\Omega})} \sum_{i=1}^{N}\sum_{j=1}^{J^{(n)}} \widehat{\omega}_i^{(n),j}\ln\varphi\left(\widehat{\boldsymbol{\alpha}}_i^{(n),j}-\boldsymbol{\alpha}\,;\boldsymbol{\Omega}\right), \quad (16)$$

yields $(\boldsymbol{\alpha}^{(n+1)}, \boldsymbol{\Omega}^{(n+1)})$, the second part of $\boldsymbol{\theta}^{(\mathrm{n}+1)}$. 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 Supplemental Appendix 10.2 for more details on the estimation procedure).

9.2. Ex-post distribution of random parameters

Estimation standard errors as well as mean and standard deviation of the ex-post distribution of random parameters η_c, μ_{cl}, ρ_0 and ρ are gathered in Table 1. Indices 1, 2 and 3 represent respectively the very-high-yielding, high-yielding and low-input CMPs.

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-economic costs (*e.g.*, environmental concerns) -- from the initial and transition probability functions. The reference CMP being the most intensive one (*i.e.* $\eta_1 = 0$), η_2 and η_3 represent the relative cost of the high-yielding and low-input technology in the initial probability function of CMP choice. On average, the high-yielding CMP is less expensive than the very high-yielding one. On the other hand, the low-input CMP appears as more expensive, suggesting learning and opportunity costs associated to the low-input CMP.
Likewise, in the transition probability the most intensive CMP type is defined as the reference one with switching costs to very-high-yield CMPs fixed at 0 ($\mu_{c1} = 0, \forall\, c \in \mathcal{C}$). 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 costly option. As for the high-yielding technology, we can observe that μ_{12} is negative, *i.e.* it is on average less expensive to switch from a very high-yielding CMP to a high-yielding one than to keep the very high-yielding CMP.

Table 1: 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

Notes: se = standard error, sd = standard deviation.

Source: Authors' calculations.

On the other hand, μ_{22} is positive, meaning that changing to the high-input technology is less expensive than keeping the high-yielding 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 CMPs. Then, one can argue that the very high-yielding CMP has larger systematic costs than the high-yielding 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. More generally, scale parameters are less dispersed than the cost parameters, *i.e.* farmers' heterogeneity is greater when considering technology costs.

The main estimates (and related estimation standard error estimates) of the expectation, variance and covariance of the parameters of the distribution of the random parameter vector of the RPHM model are presented in Table 2. Note, however, that the random parameters considered in the core of the article, that is to say the a and b terms, are defined as transformations of normal underlying random parameters, that is to say the $\tilde{a}$ and $\tilde{b}$ terms, with:

$$b^1 = \exp(\tilde{b}^1) \quad \text{and} \quad a^c = \frac{1}{1+\exp(\tilde{a}^c)}.$$

Using such transformations is a convenient means to enforce the constraints to be met by parameters *a* and *b*. Term *b* is required to be positive while a must lie in the unit interval. Table 2 provides estimates of the main parameters of the joint multivariate normal distribution of the $\tilde{a}$ and $\tilde{b}$ terms.

Table 2: Estimates and estimation standard error of the main parameters of the probability distribution of the CMP specific yield target level and inputs requirement random parameters

	Estimation	*Estimation se*
Very-high-yielding		
$E[\tilde{b}_y^1]$	2.071	0.004
$E[\tilde{b}_{x,1}^1]$	0.800	0.002
$E[\tilde{b}_{x,2}^1]$	0.655	0.006
$V[\tilde{b}_y^1]$	0.011	0.0003
$V[\tilde{b}_{x,1}^1]$	0.004	0.0001
$V[\tilde{b}_{x,2}^1]$	0.025	0.001
$cov[\tilde{b}_y^1, \tilde{b}_{x,1}^1]$	0.003	0.0001
$cov[\tilde{b}_y^1, \tilde{b}_{x,2}^1]$	0.003	0.0004
$cov[\tilde{b}_{x,1}^1, \tilde{b}_{x,2}^1]$	0.003	0.0002
High-yielding		
$E[\tilde{a}_y^2]$	3.015	0.005
$E[\tilde{a}_{x,1}^2]$	2.790	0.003
$E[\tilde{a}_{x,2}^2]$	2.236	0.006

$V[\tilde{a}_y^2]$	0.023	0.001
$V[\tilde{a}_{x,1}^2]$	0.009	0.0003
$V[\tilde{a}_{x,2}^2]$	0.031	0.001
$cov[\tilde{a}_y^2, \tilde{a}_{x,1}^2]$	0.011	0.001
$cov[\tilde{a}_y^2, \tilde{a}_{x,2}^2]$	0.024	0.001
$cov[\tilde{a}_{x,1}^2, \tilde{a}_{x,2}^2]$	0.014	0.001
Low-input		
$E[\tilde{a}_y^3]$	2.533	0.005
$E[\tilde{a}_{x,1}^3]$	2.604	0.004
$E[\tilde{a}_{x,2}^3]$	2.563	0.005
$V[\tilde{a}_y^3]$	0.018	0.001
$V[\tilde{a}_{x,1}^3]$	0.011	0.0004
$V[\tilde{a}_{x,2}^3]$	0.017	0.001
$cov[\tilde{a}_y^3, \tilde{a}_{x,1}^3]$	-0.012	0.001
$cov[\tilde{a}_y^3, \tilde{a}_{x,2}^3]$	-0.009	0.001
$cov[\tilde{a}_{x,1}^3, \tilde{a}_{x,2}^3]$	0.011	0.0004

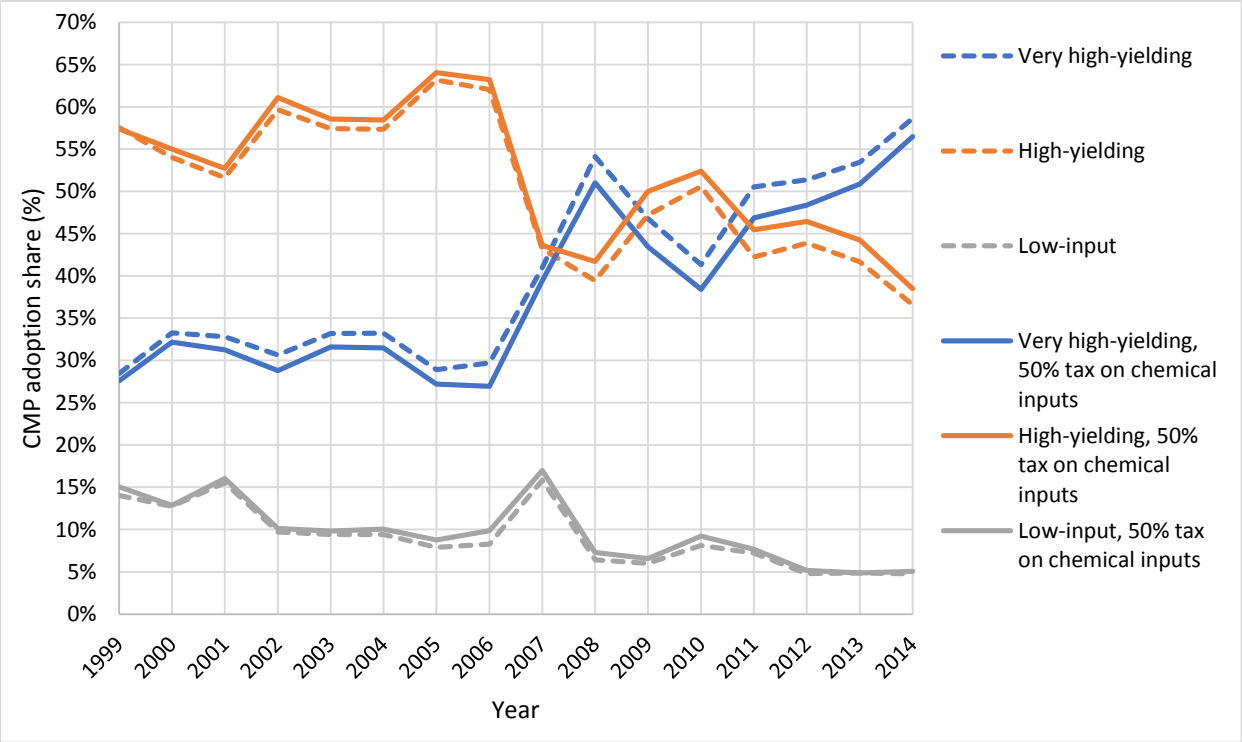
Source: Authors' calculations.

9.3. Results from chemical input tax simulations

As evoked in the core of the article, we performed simulations based on the RPHM model results to investigate into farmers' CMP adoption decision. First, we simulate a 50% and 100% tax on chemical inputs. We present in Figure 8 - respectively in Figure 9 - the results from the 50% - respectively 100% - tax on chemical inputs. As we can see, such tax mainly entails a reduction in the very high-yielding share in favor to the high-yielding one. When

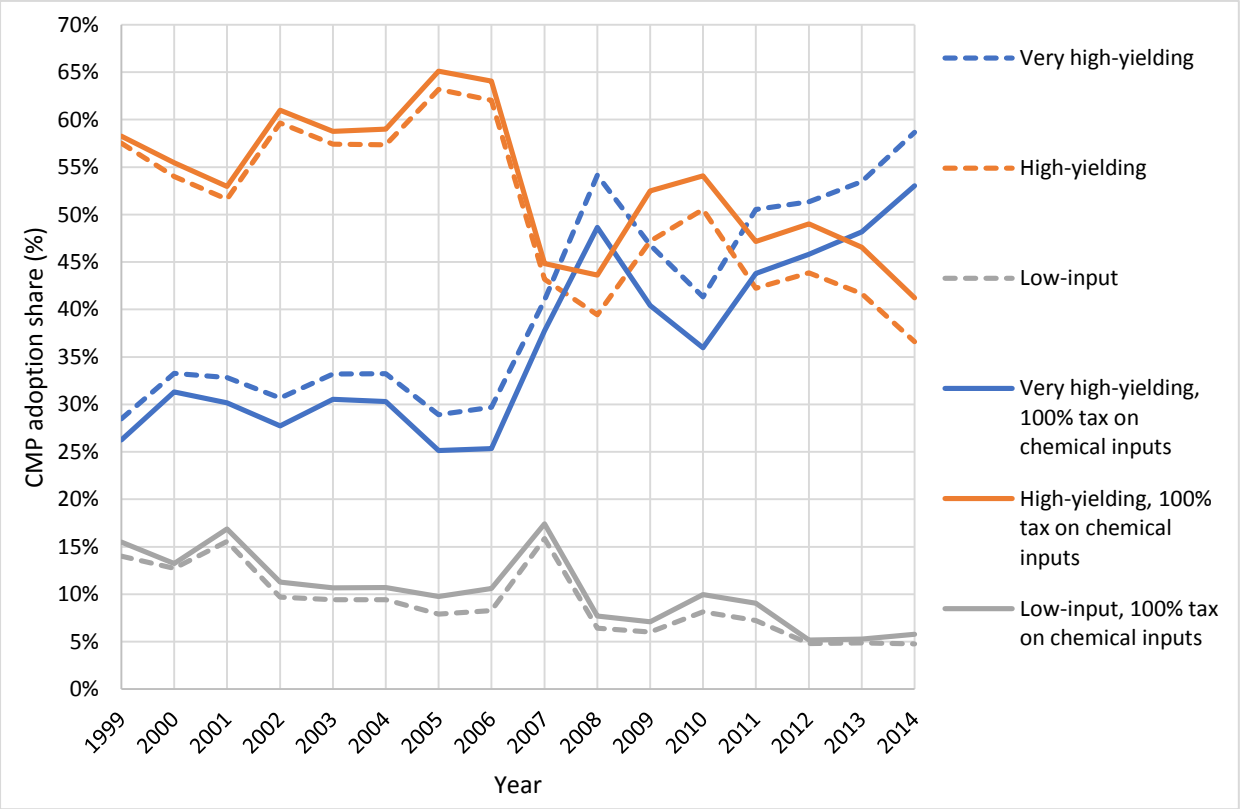
considering the 100% tax we find that, on average, the size of the very high-yielding CMP decreases by 10.25% each year while the high-yielding share increases by 6.30%. As for the low-input CMP, its share increases by 15.70% on average each year. Yet, despite this rate, the low-input CMP remains marginal among wheat producers.

Figure 8: Mean annual change in the CMP adoption share after simulating a 50% tax on chemical inputs, north eastern France, 1999–2014



Source: Authors' calculations.

Figure 9: Mean annual change in the CMP adoption share after simulating a 1000% tax on chemical inputs, north eastern France, 1999–2014



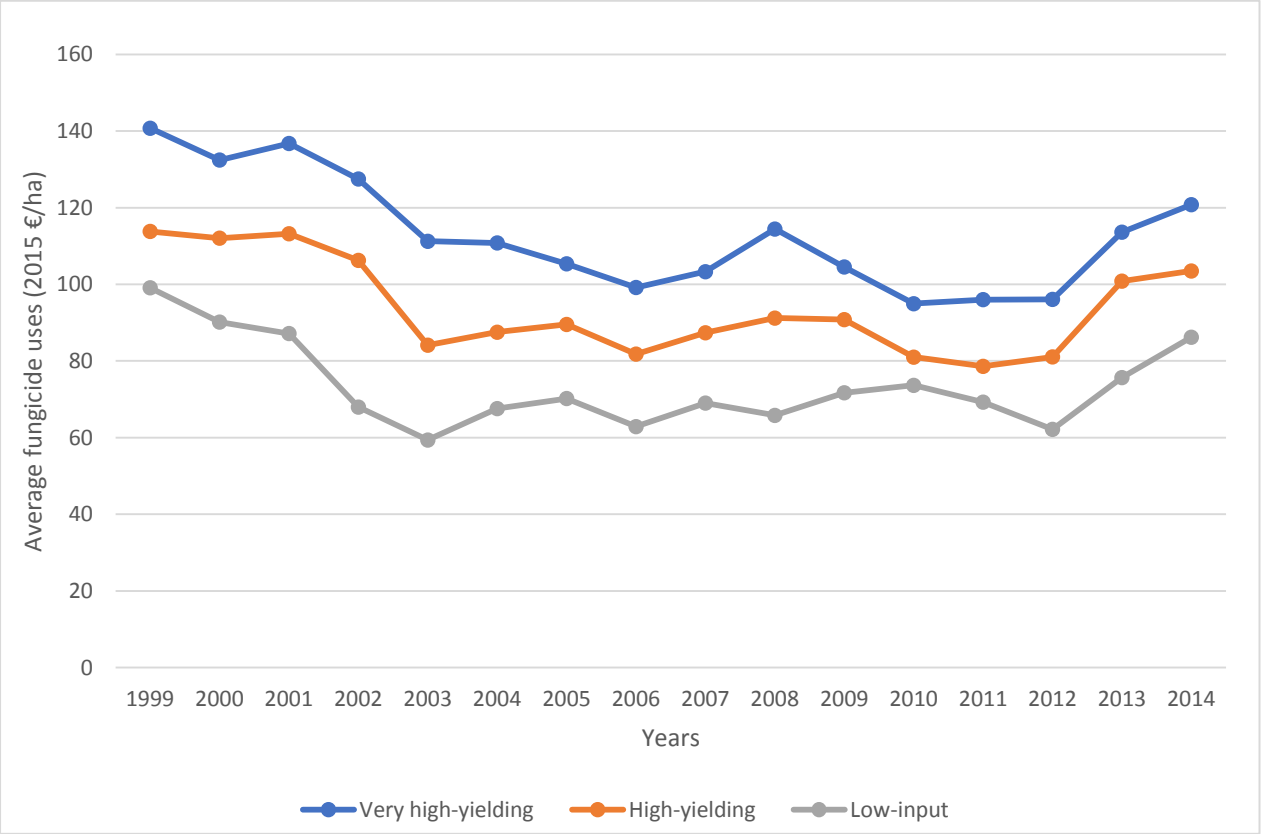
Source: Authors' calculations.

10. Supplemental Appendices

10.1. Detailed pesticide analysis

In addition to overall pesticide uses, we investigate insecticide, fungicide and herbicide uses. From Figures 10, 11 and 12, we can see that fungicide uses is the most discriminating pesticide category between our three CMP categories. As for herbicides and insecticides, the discrimination is well-marked for very high-yielding CMPs *versus* high-yielding and low-input CMPs. Yet, uses among high-yielding and low-input CMPs tend to overlap, implying a lower discrimination power of herbicides and insecticides than fungicides. A closer look at the relative differences between very high-yielding and low-input fungicide, herbicide and insecticide uses.

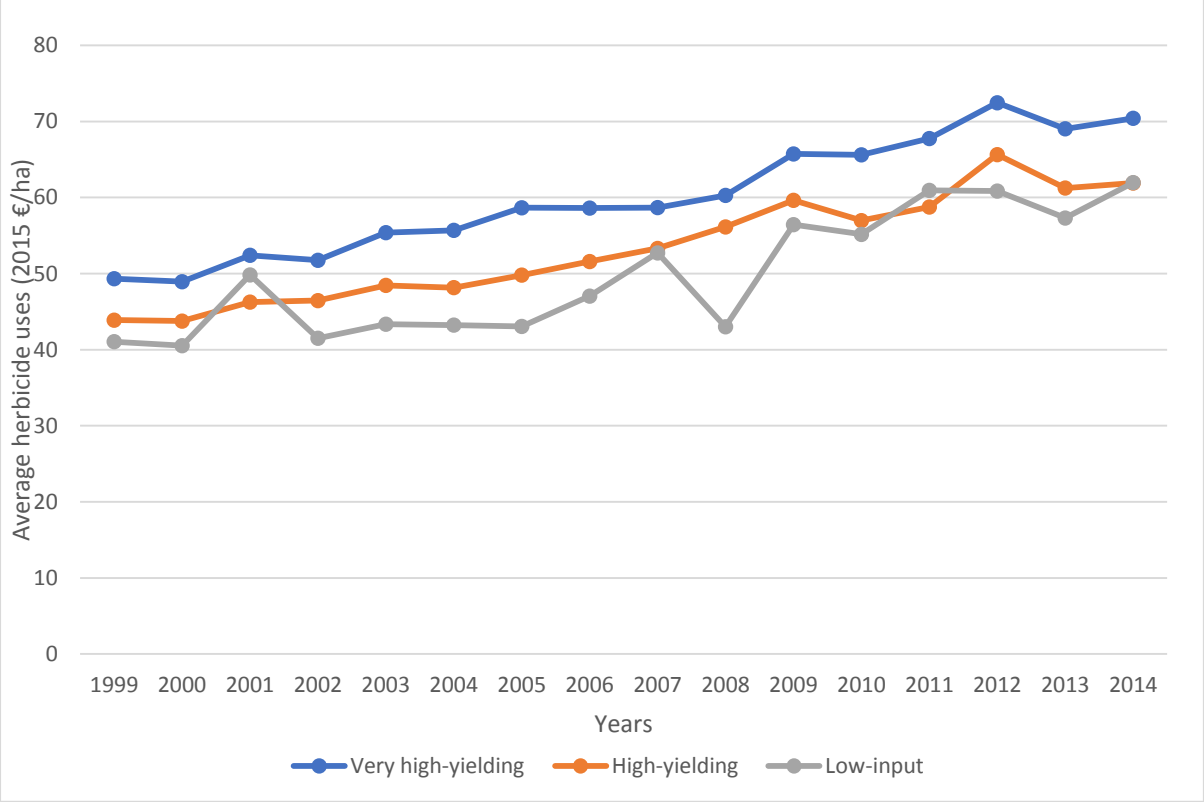
Figure 10: Annual mean fungicide use per CMP type, north eastern France, 1999–2014



Source: Authors' calculations.

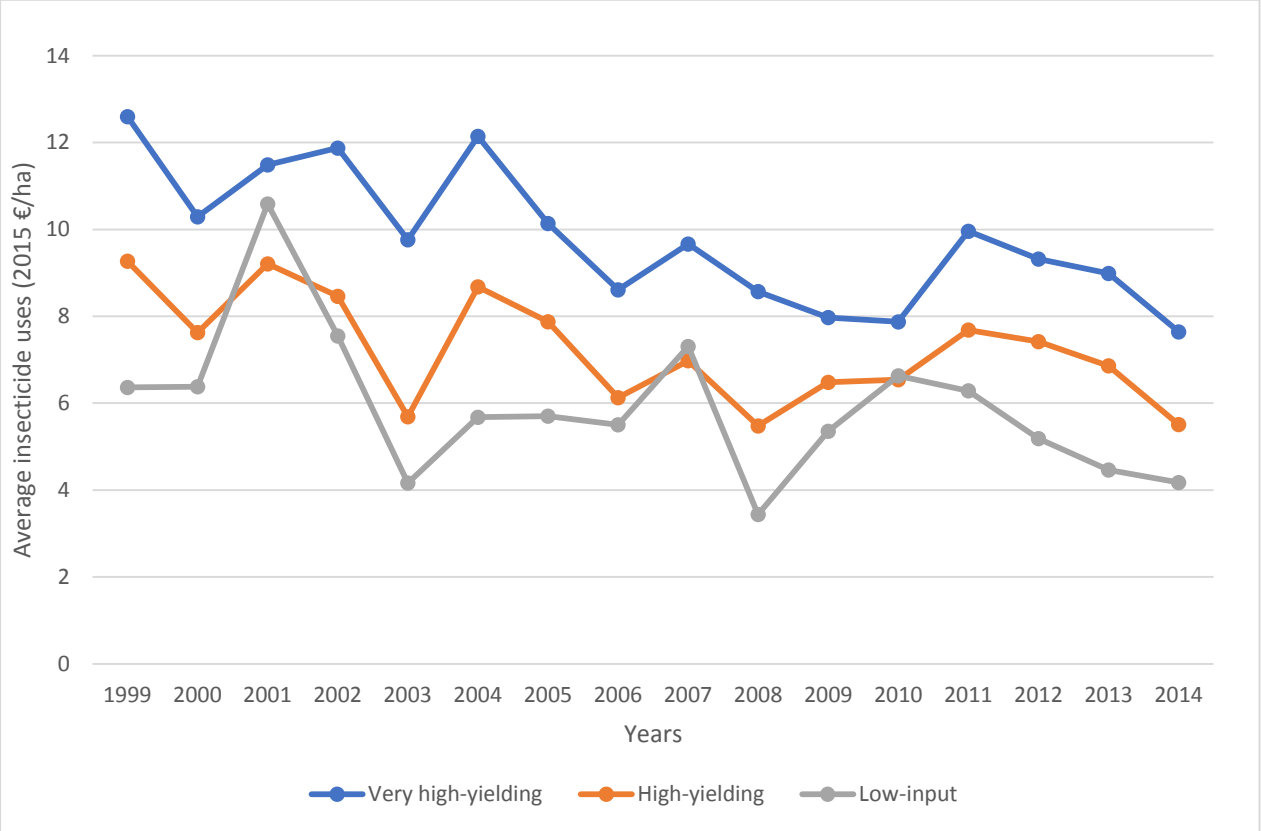
On average, fungicide and herbicide uses are 33% lower in low-input CMPs than in very high-yielding ones. When targeting lower yields, farmers can use more resistant crop varieties and lower their sowing density, hence reducing their need in fungicides. As for herbicides, they can be – at least partially – substituted for by mechanical weeding. Differences in insecticide uses are smaller as on average, low-input CMPs use 21% less insecticide than very high-yielding ones. Even if pest infestations can be partially avoided by a well-designed crop rotation scheme and more resistant crop varieties they cannot be totally avoided. And insecticides remain the most efficient way to get rid of them. This could explain why the difference is less marked for insecticide uses compared to herbicide and fungicide uses.

Figure 11: Annual mean herbicide use per CMP type, north eastern France, 1999–2014



Source: Authors' calculations.

Figure 12: Annual mean insecticide use per CMP type, north eastern France, 1999–2014



Source: Authors' calculations.

10.2. SAEM algorithm explicit forms

Terms $E^{(n)}\left[ln\,\ell_i^C(\boldsymbol{\theta}\,|\,\boldsymbol{r}_{(i)},\boldsymbol{\alpha}_i)|\boldsymbol{\zeta}_i,\boldsymbol{\alpha}_i\right]$ 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 $\boldsymbol{\theta}$, $\boldsymbol{\theta}^{(n-1)}$:

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

$$s_{it,c}^{(\zeta,n)} = s_{it,c}^{(\zeta,n-1)} + \lambda_{(n)}\left(R^{-1}\sum_{m=1}^{R} p_t\left(c \,|\mathbf{w}_{(i)}, \widehat{\boldsymbol{\alpha}}_i^{(n,m)}\right) - s_{it,c}^{(\zeta,n-1)}\right)$$

$$s_{it,cl}^{(\xi,n)} = s_{it,cl}^{(\xi,n-1)} + \lambda_{(n)}\left(R^{-1}\sum_{m=1}^{R} s_t\left(c, l \,|\mathbf{w}_{(i)}, \widehat{\boldsymbol{\alpha}}_i^{(n,m)}\right) - s_{it,cl}^{(\xi,n-1)}\right)$$

$$\boldsymbol{s}_i^{(\boldsymbol{\alpha},n)} = \boldsymbol{s}_i^{(\boldsymbol{\alpha},n-1)} + \lambda_{(n)}\left(R^{-1}\sum_{m=1}^{R} \widehat{\boldsymbol{\alpha}}_i^{(n,m)} - \boldsymbol{s}_i^{(\boldsymbol{\alpha},n-1)}\right)$$

$$\boldsymbol{s}_i^{(\boldsymbol{\alpha}\boldsymbol{\alpha},n)} = \boldsymbol{s}_i^{(\boldsymbol{\alpha}\boldsymbol{\alpha},n-1)} + \lambda_{(n)}\left(R^{-1}\sum_{i=1}^{N}\sum_{m=1}^{R} \widehat{\boldsymbol{\alpha}}_i^{(n,m)}\left(\widehat{\boldsymbol{\alpha}}_i^{(n,m)}\right)' - \boldsymbol{s}_i^{(\boldsymbol{\alpha}\boldsymbol{\alpha},n-1)}\right)$$

$$\boldsymbol{s}_{it,c}^{(\boldsymbol{\varepsilon},n)} = \boldsymbol{s}_{it,c}^{(\boldsymbol{\varepsilon},n-1)} + \lambda_{(n)}\left(R^{-1}\sum_{m=1}^{R} p_t\left(c \,|\mathbf{w}_{(i)}, \widehat{\boldsymbol{\alpha}}_i^{(n,m)}\right)\left(\mathbf{q}_{it} - \mathbf{b}_i^{c,(n,m)}\right) - \boldsymbol{s}_{it,c}^{(\boldsymbol{\varepsilon},n-1)}\right)$$

$$\boldsymbol{s}_c^{(\boldsymbol{\varepsilon}\boldsymbol{\varepsilon},n)} = \boldsymbol{s}_c^{(\boldsymbol{\varepsilon}\boldsymbol{\varepsilon},n-1)} + \lambda_{(n)}\left(R^{-1}\sum_{i=1}^{N}\sum_{t=t_{(i)}}^{T_{(i)}}\sum_{m=1}^{R} p_t\left(c \,|\mathbf{w}_{(i)}, \widehat{\boldsymbol{\alpha}}_i^{(n,m)}\right)\left(\mathbf{q}_{it} - \mathbf{b}_i^{c,(n,m)}\right)\left(\mathbf{q}_{it} - \mathbf{b}_i^{c,(n,m)}\right)' - \boldsymbol{s}_c^{(\boldsymbol{\varepsilon}\boldsymbol{\varepsilon},n-1)}\right)$$

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

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

$$\approx$$

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

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

$$\boldsymbol{\alpha}_0^{(n)}=\;N^{-1}\sum_{i=1}^{N}\boldsymbol{s}_i^{(\boldsymbol{\alpha},n)}$$

$$\boldsymbol{\Omega}^{(\boldsymbol{n})}=N^{-1}\boldsymbol{s}_i^{(\boldsymbol{\alpha\alpha},n)}-\boldsymbol{\alpha}_0^{(n)}\left(\boldsymbol{\alpha}_0^{(n)}\right)'$$

$$\boldsymbol{\delta}^{(\boldsymbol{n})}=\left(\sum_{i=1}^{N}\sum_{\boldsymbol{t}=t_{(i)}}^{T_{(i)}}\sum_{c\in\mathcal{C}}s_{it,c}^{(\boldsymbol{\zeta},n)}\,\boldsymbol{z}_{it}'\left(\boldsymbol{\Sigma}^{c,(n-1)}\right)^{-1}\mathbf{z}_{it}\right)^{-\mathbf{1}}\left(\sum_{i=1}^{N}\sum_{\boldsymbol{t}=t_{(i)}}^{T_{(i)}}\sum_{c\in\mathcal{C}}\boldsymbol{z}_{it}'\left(\boldsymbol{\Sigma}^{c,(n-1)}\right)^{-1}\boldsymbol{s}_{it,c}^{(\boldsymbol{\varepsilon},n)}\right)$$

$$\boldsymbol{\Sigma}^{c,(n)}=\left(\sum_{i=1}^{N}\sum_{\boldsymbol{t}=t_{(i)}}^{T_{(i)}}s_{it,c}^{(\boldsymbol{\zeta},n)}\right)^{-1}\left(\boldsymbol{s}_c^{(\boldsymbol{\varepsilon\varepsilon},n)}-\sum_{i=1}^{N}\sum_{\boldsymbol{t}=t_{(i)}}^{T_{(i)}}\mathbf{z}_{it}\boldsymbol{\delta}^{(\boldsymbol{n})}\left(s_{it,c}^{(\boldsymbol{\zeta},n)}\right)'\right.$$

$$\left.-\sum_{i=1}^{N}\sum_{\boldsymbol{t}=t_{(i)}}^{T_{(i)}}\left(\mathbf{z}_{it}\boldsymbol{\delta}^{(\boldsymbol{n})}\left(s_{it,c}^{(\boldsymbol{\zeta},n)}\right)'\right)'+\sum_{i=1}^{N}\sum_{\boldsymbol{t}=t_{(i)}}^{T_{(i)}}s_{it,c}^{(\boldsymbol{\zeta},n)}\mathbf{z}_{it}\boldsymbol{\delta}^{(\boldsymbol{n})}\left(\mathbf{z}_{it}\boldsymbol{\delta}^{(\boldsymbol{n})}\right)'\right)$$

a. 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 maximum likelihood. Kuhn and Lavielle (2005) proposes to set $\lambda_{(n)} = 1$ for the first n_1 iterations and then gradually reduce $\lambda_{(n)}$. We set here:

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

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

b. Simulation step procedure

To perform (S) step at iteration n, we use a few Markov chain Monte-Carlo (MCMC) iterations with $p\left(\boldsymbol{\alpha}|\boldsymbol{\zeta}_i, \boldsymbol{\theta}^{(n-1)}\right)$ as stationary distribution, and we retain R MCMC draws for each $i, i = 1, \ldots, N$.[13] 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 $\widehat{\boldsymbol{\alpha}}_{i,m}$ such that $\widehat{\boldsymbol{\alpha}}_{i,m} \sim \mathcal{N}(\widehat{\boldsymbol{\alpha}}_{i,m-1}, \boldsymbol{\Psi})$ for $1 \leq i \leq N$ and $1 \leq m \leq R + R_{\text{burn}}$. We defined the acceptance rate as:

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

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

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

c. Stopping rule of the algorithm and diagnostic plots

As in Koutchadé et al. (2018, 2021) we use a standard stopping rule based on the relative changes in the values of the estimated parameters between two iterations (Booth et al., 2001; Booth and Hobert, 1999). If the condition

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

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.

d. 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 $\{\widehat{\boldsymbol{\alpha}}_{i,r} : r = 1, \dots, R\}$ from $p(\boldsymbol{\alpha}|\boldsymbol{\zeta}_i, \widehat{\boldsymbol{\theta}})$, for $i = 1, \dots, N$, where $\widehat{\boldsymbol{\theta}}$ are the estimates we obtained from the SAEM algorithm. Then, we can approximate the information matrix $I(\widehat{\boldsymbol{\theta}})$ by:

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

and the variance of estimates by:

$$V(\widehat{\boldsymbol{\theta}}) = \widetilde{I}(\widehat{\boldsymbol{\theta}})^{-1}.$$

e. Estimation of the likelihood and model selection

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

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

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

$p(\mathbf{q}_{(i)}, \widehat{\boldsymbol{\theta}})$ has no closed form, so we use the importance-sampling approach to estimate it. From prior distribution $\varphi(\boldsymbol{\alpha}_i - \widehat{\boldsymbol{\alpha}};\ \widehat{\boldsymbol{\Omega}})$ as importance density, we draw independence sequence $\{\boldsymbol{\alpha}_{i,r} : r = 1, \dots, R\}$ and then approximate $p(\mathbf{q}_{(i)}, \widehat{\boldsymbol{\theta}})$ by

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

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

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

$$\begin{cases} -2LL = -2\ell(\widehat{\boldsymbol{\theta}}) \\ AIC = 2\ell(\widehat{\boldsymbol{\theta}}) + 2P \\ BIC = 2\ell(\widehat{\boldsymbol{\theta}}) + ln(N)P \end{cases}$$

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

f. Forward-Backward algorithm

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

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

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

We can show that (Maruotti, 2011):

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

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

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

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

$$\beta_{it}(c, \boldsymbol{\alpha}_i) = p\left(q_{it+1}, , \dots, q_{iT_{(i)}} \mid r_{it} = c, \boldsymbol{\alpha}_i\right),$$

for $i = 1, \dots, N$, $t = t_{(i)}, \dots, T_{(i)}$ and $c \in \mathcal{C}$. $\beta_{it}(c, \boldsymbol{\alpha}_i)$ denotes the probability of the partial sequence $\left(\mathbf{q}_{it+1}, \dots, \mathbf{q}_{iT_{(i)}}\right)$ given that farmer i chooses CMP c at time t and given the random parameter α_i. We can compute this term iteratively by:

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

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

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

We also show that:

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

g. Estimation of the individual parameters and sequences of CMP

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

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

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

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

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

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

h. Viterbi algorithm

Let start by defining the following Viterbi path probability:

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

This term can be computed iteratively by:

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

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

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