Crop choices in micro-econometric multi-crop models: modelling corners, kinks and jumps

Obafemi Philippe Koutchade, Alain Carpentier, Fabienne Femenia

Working Papers SMART - LERECO

Crop choices in micro-econometric multi-crop models: modelling corners, kinks and jumps

Obafèmi Philippe KOUTCHADÉ, Alain CARPENTIER, Fabienne FÉMÉNIA

Working Paper SMART – LERECO N°20-09

October 2020

Paper published in the *American Journal of Agricultural Economics*
DOI: https://doi.org/10.1111/ajae.12152

UMR INRAE-Agrocampus Ouest **SMART - LERECO**
(Laboratoires d'Etudes et de Recherches en Economie sur les Structures et Marchés Agricoles, Ressources et Territoires)

Les Working Papers SMART-LERECO ont pour vocation de diffuser les recherches conduites au sein des unités SMART et LERECO dans une forme préliminaire permettant la discussion et avant publication définitive. Selon les cas, il s'agit de travaux qui ont été acceptés ou ont déjà fait l'objet d'une présentation lors d'une conférence scientifique nationale ou internationale, qui ont été soumis pour publication dans une revue académique à comité de lecture, ou encore qui constituent un chapitre d'ouvrage académique. Bien que non revus par les pairs, chaque working paper a fait l'objet d'une relecture interne par un des scientifiques de SMART ou du LERECO et par l'un des deux éditeurs de la série. Les Working Papers SMART-LERECO n'engagent cependant que leurs auteurs.

The SMART-LERECO Working Papers are meant to promote discussion by disseminating the research of the SMART and LERECO members in a preliminary form and before their final publication. They may be papers which have been accepted or already presented in a national or international scientific conference, articles which have been submitted to a peer-reviewed academic journal, or chapters of an academic book. While not peer-reviewed, each of them has been read over by one of the scientists of SMART or LERECO and by one of the two editors of the series. However, the views expressed in the SMART-LERECO Working Papers are solely those of their authors.

Crop choices in micro-econometric multi-crop models: modelling corners, kinks and jumps

Obafèmi Philippe KOUTCHADÉ

INRAE, UMR1302 SMART-LERECO, 35000, Rennes, France

Alain CARPENTIER

INRAE, UMR1302 SMART-LERECO, 35000, Rennes, France

Fabienne FEMENIA

INRAE, UMR1302 SMART-LERECO, 35000, Rennes, France

Paper published in the *American Journal of Agricultural Economics*

DOI: https://doi.org/10.1111/ajae.12152

Acknowledgements

This research has received funding from Arvalis - Institut du végétal and has benefited from financial support by the MINDSTEP project, funded by the European Union's Horizon 2020 research and innovation programme under grant agreement No 817566.

Corresponding author

Fabienne Femenia
UMR SMART-LERECO, INRAE,
4, Allée Adolphe Bobierre
35011 Rennes, France
Email : Fabienne.Femenia@rennes.inra.fr
Téléphone / Phone : +33 (0)2 23 48 56 10
Fax : +33 (0)2 23 48 53 80

Crop choices in micro-econometric multi-crop models: modelling corners, kinks and jumps

Abstract

Null crop acreages raise pervasive issues when modelling acreage choices with farm data. We revisit these issues and emphasize that null acreage choices arise not only due to binding non-negativity constraints but also due to crop production fixed costs. Based on this micro-economic background, we present a micro-econometric multi-crop model that consistently handles null acreages and accounts for crop production fixed costs. This multivariate endogenous regime switching model allows for specific crop acreage patterns, such as multiple kinks and jumps in crop acreage responses to economic incentives, that are due to changes in produced crop sets. Currently available micro-econometric multi-crop models, which handle null acreages based on a censored regression approach, cannot represent these patterns.

We illustrate the empirical tractability of our modelling framework by estimating a random parameter version of the proposed endogenous regime switching micro-econometric multi-crop model with a panel dataset of French farmers. Our estimation and simulation results support our theoretical analysis, the effects of crop fixed costs and crop set choices on crop acreage choices in particular. More generally, these results suggest that the micro-econometric multi-crop model presented in this article can significantly improves empirical analyses of crop supply based on farm data.

Keywords: acreage choice, crop choice, endogenous regime switching, random parameter models

JEL classifications: Q12, C13, C15

Choix des cultures et solutions en coin dans les modèles micro-économétriques de choix de production multiculture

Résumé

Les surfaces de culture nulles soulèvent de nombreuses questions lorsque l'on souhaite modéliser les choix d'assolement des agriculteurs à partir de données individuelles. Nous revisitons ces questions en mettant en avant le fait que les choix de surfaces nulles ne sont pas uniquement dus à des contraintes de non-négativité saturées mais également à des coûts fixes de production des cultures. A partir de ce cadre micro-économique, nous présentons un modèle micro-économétrique multiculture permettant de représenter de façon cohérente les choix surfaces nulles et tenant compte des coûts fixes de production des cultures. Ce modèle multivarié à changement de régime endogène offre la possibilité de représenter des schémas de choix d'assolement spécifiques, présentant notamment de multiples sauts et points d'inflexion dans les réponses des assolements aux incitations économiques en raison de changements dans les ensembles de cultures produites. Les modèles économétriques multicultures disponibles actuellement, qui tiennent compte des surfaces nulles à partir d'approches basées sur des régressions censurées, ne permettent pas de représenter ce type de schéma.

Nous illustrons la tractabilité empirique de notre approche en estimant, sur un panel d'agriculteurs français, une version à paramètres aléatoires du modèle micro-économétrique multiculture à changement de régime endogène que nous proposons. Nos résultats d'estimations et de simulations viennent appuyer notre analyse théorique, concernant en particulier les effets des coûts fixes de production des cultures et des choix d'ensembles de cultures produites sur les choix d'assolement. Plus généralement, le modèle micro-économétrique multiculture présenté ici peut améliorer significativement les analyses empiriques d'offre de culture basées sur des données individuelles d'exploitation.

Mots-clés : choix d'assolement, choix de culture, changement de régime endogène, modèle à paramètres aléatoires

Classifications JEL: Q12, C13, C15

Choix des cultures et solutions en coin dans les modèles micro-économétriques de choix de production multiculture

1. Introduction

Market prices and agricultural policies impact crop supplies through their effects on input uses and yield levels, and acreage choices. Starting in the eighties with the pioneering work of Just *et al.* (1983), Chambers and Just (1989) and Chavas and Holt (1990), agricultural production economists developed micro-econometric multi-crop (MEMC) models for analyzing and quantifying these effects with farm accountancy data. These models have then been widely applied during the last decades.

In MEMC models, farmers are assumed to allocate their cropland to the crops of a given crop set in order to maximize their expected profit or the expected utility of their profit. This ensures the economic consistency of the resulting models. However, currently available MEMC models ignore or poorly describe an important decision of crop producers: their choice to produce a subset of crops among the set of crops they can produce and sell. Indeed, applications of MEMC models frequently ignore null acreages by relying either on very specific farm samples (*e.g.*, Just *et al.*, 1983, 1990; Bayramoglu and Chakir, 2016) or on crop aggregation that eliminate null crop acreages (*e.g.*, Oude Lansing and Peerlings, 1996; Serra *et al.*, 2005; Oude Lansink, 2008; Carpentier and Letort, 2012, 2014).[1] Yet, sample selection prevents extrapolation of the estimation results to farmers not producing all considered crops while crop aggregation induces information losses regarding production choices at the crop level.

A few recent MEMC models explicitly account for null crop acreages (*e.g.*, Sckokai and Moro, 2006, 2009; Lacroix and Thomas, 2011; Bateman and Fezzi, 2011; Platoni *et al.*, 2012).[2] These models are designed as censored regression (CR) systems and are estimated following two-step approaches inspired by that initially proposed by Shonkwiler and Yen (1999). These MCEM based on censored regressions (CR-MCEM) suitably account for null acreages from a statistical viewpoint but display severe micro-economic inconsistencies.

---

[1] Other applications rely on aggregated data in which the occurrence of null acreages is limited (*e.g.*, Chavas and Holt, 1990; Moore and Negri, 1992; Coyle, 1992, 1999; Ozarem and Miranowski, 1994; Guyomard *et al*, 1996). Such applications are now less frequent thanks to increasing availability of micro-data.

[2] The early studies by Moore and Negri (1992) and Moore *et al* (1994) being notable exceptions in this respect.

Indeed, these models are conceived as statistical versions – featuring error terms and accounting for mass points at 0 – of theoretical micro-economic models ignoring null acreages. Their main shortcoming is due to their relying on a single crop acreage choice model, whatever the subset of crops actually produced. These models thus fail to recognize that the crop acreage choices of a farmer structurally depend on the composition of the set of crops actually produced by this farmer. For instance, farmers are unlikely to consider the prices of the crops they don't produce when choosing the acreages of the crops they produce.[3] Of course, the lack of micro-economic coherency of CR-MCEM models substantially undermines their ability to yield consistent estimates of crop acreage responses to economic incentives. The composition of the produced crop sets, which displays substantial variability when null acreages are frequent, deeply impacts the structure of farmers' crop acreage choices. These effects of farmers' crop set choices are ignored in CR-MCEM models.

The recent articles addressing the issue raised by null crop acreages from a statistical viewpoint by considering CR-MEMC models don't focus either on corner solutions in acreage choices or on farmers' crop choices. By contrast, the main objective of this article is to develop a consistent modelling framework for analyzing farmers' crop set choices and, as a result, for handling null crop acreages in MEMC models. More precisely, the main aims of this article are (a) to revisit the null acreage issue in multi-crop models from a theoretical viewpoint, (b) to propose an original MEMC model that accounts for farmers' crop choices in a way that is consistent from an economic viewpoint, together with a suitable estimation approach, and (c) to show, by means of an empirical application focused on crop diversification choices, that considering crop set choices significantly enriches micro-econometric analyses of farmers' crop supply.

Our multi-crop micro-economic modelling framework is based on an expected profit maximization problem considering land as an allocable quasi-fixed input. This problem includes the usual crop acreage non-negativity constraints but also production regime fixed costs. The production regime chosen by a farmer is defined by the subset of crops that this farmer decides to plant. The regime fixed costs consist of unobserved costs – such as unmeasured marketing costs or implicit labor and machinery management costs – that depend on the set of crops that are grown but that don't depend on the acreages of these crops.

[3] They don't in a static framework although they might in a dynamic one. For instance, forward looking farmers consider (future) prices of crops they don't currently produce if they plan to produce these crops in the future and if their current crop acreage choices impact their future expected profits (*e.g.*, due to crop rotation effects). Yet, if they cannot be ruled out, these (indirect) effects of the non-produced crop prices are likely to be limited.

Accordingly, our modelling framework assumes that farmers choose the production regime maximizing their expected profit, regime fixed costs included, as well as the related optimal crop acreage, yield and input use levels. Importantly, our considering regime fixed costs implies that null acreages are not necessarily due to binding non-negativity constraints.

Based on this micro-economic background, we design our MEMC model as an endogenous regime switching (ERS) multivariate model with multiple regimes. This endogenous regime switching micro-econometric multi-crop (ERS-MEMC) model consists of a probabilistic regime choice model coupled with a set of regime specific MEMC models. As estimating multivariate ERS models with multiple regimes is challenging and considering regime specific fixed costs increases the estimation burden, choosing relevant functional forms for the per regime MEMC models appears crucial. Thanks to their specific properties, the Multinomial Logit (MNL) acreage choice models proposed by Carpentier and Letort (2014) are particularly well suited in that respect. They yield simple and well-behaved functional forms for important components of our ERS-MEMC model, thereby significantly reducing its estimation cost. Relying on the MNL acreage choice models also enables us, following Koutchadé *et al.* (2018), to go one step further and to account for farmers' unobserved heterogeneity by considering a random parameter version of our ERS-MEMC model. Estimating ERS models with multiple regimes is challenging mostly because their likelihood function involves integration of expectations over the probability distribution of multivariate latent error terms (*e.g.*, Pudney, 1989).[4] Also, the likelihood function of our ERS-MEMC model needs to be integrated over the probability distribution of its random parameters. Our estimation approach combines tools from the micro-econometrics and computational statistics literatures.

We illustrate the empirical tractability of our approach by estimating our model for a panel data sample of French arable crop producers. Our results tend to demonstrate that our random parameter ERS-MEMC model performs well according to standard fit criteria. They also tend to show that regime specific fixed costs significantly matter in farmers' crop choice, along with crop expected returns. Importantly, these results also demonstrate that acreage choices' responses to economic incentives strongly depend on the production regime choices. The elasticity of crop acreages in crop prices increases in the number of produced crops, a pattern that cannot be reproduced by CR-MEMC models. Finally, our simulation results show that the

[4] In particular, the probability function of the production regime choice of our model cannot be integrated analytically. We also have to overcome the fact that the crop yield and input use (and, thus, expected crop return) levels of the crops that are not produced are unobserved.

acreage of minor crops respond non-linearly to increases in their prices due to production regime changes.

Our contributions are twofold. First, ERS-MEMC model presented in this article accounts for null crop acreages while relying on a well-defined micro-economic background. As a result, it is the first theoretically coherent response to an issue that is pervasive when analyzing crop production with farm level data. Other ERS-MEMC models could be considered, but the one presented here allows to consider production regime fixed costs as well as farm specific parameters while remaining empirically tractable. Second, this model allows to disentangle the effects of the main economic drivers of farmers' crop supply choices. It accounts for intensive and extensive margin choices, including the effects on crop set choices at the extensive margin. This unique feature is of special interest for investigating future agri-environmental policies. In particular, owing to its positive agronomic effects, crop diversification is a key feature of environmentally friendly crop production systems (*e.g.*, Matson *et al.*, 1997; Tilman *et al.*, 2002; Lin, 2011; Kremen *et al.*, 2012; Bowman and Zilberman, 2013). Our modelling framework is especially well-suited for analyzing samples containing both specialized and diversified farms as well as for simulating the effects of policy instruments aimed to foster crop diversification.

The rest of this article is organized as follows. The approach proposed to account for crop choices in micro-economic models of acreage decisions is presented in the first section. The structure of the corresponding ERS-MEMC model is described in the second section. The main features of our estimation strategy are presented in the third section, with a specific focus on the main issues arising with random parameter ERS-MEMC models.[5] Illustrative estimation and simulation results are provided in the fourth section. Finally, we conclude.

2. Regime switching in multi-crop acreage models: corners, kinks and jumps

This section presents the theoretical modelling framework we propose for dealing with null crop acreages in micro-econometric acreage choice models. We proceed in three steps. First, we present the micro-economic crop acreage choice model underlying our ERS-MEMC model. Second, we compare this model to the models that have been proposed for modelling multiple

[5] The overall structure of our estimation procedure is described in the Appendix. A detailed description is given in a dedicated Online Appendix.

binding non-negativity constraints or regime switching. We focus on the ability of these models to cope with corners, kinks and jumps in farmers acreage choices.[6] Third, we present the functional form of the crop acreage choice models used in our ERS-MEMC model.

2.1. Crop choices and crop acreages

We assume that farmers can allocate their fixed cropland area to *K* crops. Accordingly, set $\mathcal{K}=\{1,...,K\}$ denotes the set of crops that any considered farmer can produce and sell and farmers' problem consists of optimally choosing a crop acreage share vector $\mathbf{s}=(s_k:k\in\mathcal{K})$ satisfying $\mathbf{s}\geq\mathbf{0}$ and $\mathbf{s}'\boldsymbol{\iota}=1$, term $\boldsymbol{\iota}$ being the dimension *K* unitary column vector.

We now introduce notions and notations aimed at describing farmers' decisions to produce a subset of crops among crop set $\mathcal{K}$. Set $\mathcal{R}=\{1,...,R\}$ denotes the set of feasible production regimes. A production regime is defined by the subset of crops with strictly positive acreages. Set $\mathcal{K}^{+}(r)$ denotes the subset of crops planted in regime *r* while $\mathcal{K}^{0}(r)$ denotes its complement to $\mathcal{K}$, that is to say the subset of crops that are not planted in regime *r*. Finally, function $\rho(\mathbf{s})$ defines the regime of the acreage share vector **s**.

We assume that farmers are risk neutral. In year *t* farmer *i* is assumed to choose her/his crop acreages by solving the following expected profit maximization problem:

$$\max_{\mathbf{s}}\left\{\mathbf{s}'\boldsymbol{\pi}_{it}-C_{it}(\mathbf{s})-D_{it}\left(\rho(\mathbf{s})\right)\;\text{ s.t. }\;\mathbf{s}\geq\mathbf{0}\;\text{ and }\;\mathbf{s}'\boldsymbol{\iota}=1\right\}. \tag{1}$$

Term $\boldsymbol{\pi}_{it}=(\pi_{k,it}:k\in\mathcal{K})$ is the vector of crop returns expected by farmer *i* when choosing **s** in year *t*. Function $C_{it}(\mathbf{s})$ is the implicit management cost of acreage **s** and $D_{it}(r)$ is the fixed cost of production regime *r* incurred by farmer *i* in *t*. This cost is fixed in the sense that it doesn't depend on **s**.

Acreage management costs $C_{it}(\mathbf{s})$ are costs not included in the crop gross margins that vary in **s**. They include unobserved variable input costs. They also account for the implicit costs related to constraints on acreage choices due to limiting quantities of machinery or labor, or to agronomic factors. These constraints providing motives for diversifying crop acreages, function $C_{it}(\mathbf{s})$ is assumed to be convex in **s**. In order to ensure that the solution in **s** to problem (1) is

[6] According to a slight adaptation of terminology introduced by Pudney (1989).

unique, we strengthen this assumption by assuming that function $C_{it}(\mathbf{s})$ is strictly convex in $\mathbf{s}$.[7] These crop acreage management costs prevent farmers to solely produce the most profitable crop.

Regime fixed cost terms $D_{it}(r)$ introduce discrete elements, and thus severe discontinuities, in farmers' acreage choices. These costs do not depend on the chosen acreage in a given regime, they only depend on the crop set defining this regime. They account for the hidden fixed costs incurred by the farmer for any acreage choice in the considered regime, such as fixed costs related to the marketing process of the crop products or those incurred when purchasing specific variable inputs, when renting specific machines, when seeking crop specific advises, *etc*. These regime fixed costs may also depend on characteristics of crop biological cycles. For instance, part-time farmers may decide not to produce a given crop because the management of this crop is not compatible with their other non-farming activities.

The smooth acreage management cost function $C_{it}(\mathbf{s})$ and the discontinuous regime fixed cost function $D_{it}\left(\rho(\mathbf{s})\right)$ are expected to impact farmers' crop diversification in opposite directions. While limiting quantities of quasi-fixed factors impose constraints fostering crop diversification, regime fixed costs are expected to foster crop specialization. In particular, the regime fixed costs are expected to be non-decreasing in the number of produced crops.[8]

We solve farmers' expected profit maximization problem following a standard backward induction approach according to which farmers choose their production regime after examining their expected profit in each possible production regime.

First, the acreage choice problem is solved for each potential regime. This yields the regime specific optimal acreage shares:

$$\mathbf{s}_{it}(r) = \arg\max_{\mathbf{s}} \left\{ \mathbf{s}'\boldsymbol{\pi}_{it} - C_{it}(\mathbf{s}) \text{ s.t. } \mathbf{s} \geq \mathbf{0},\ \mathbf{s}'\boldsymbol{\iota} = 1 \text{ and } s_k = 0 \text{ if } k \in \mathcal{K}^0(r) \right\} \tag{2a}$$

and the regime specific optimal expected profit levels (regime specific fixed costs excluded):

---

[7] Analogous cost functions are used in the Positive Mathematical Programming literature (*e.g.*, Mérel and Howitt, 2014; Heckelei *et al*, 2012) and in the multi-crop econometric literature (*e.g.*, Heckeleï and Wolff, 2003; Carpentier and Letort, 2012, 2014).

[8] Note however that in specific empirical settings the $D_{it}(r)$ terms may also capture the effects of exogenous factors preventing farmer i to produce specific crops, *e.g.* due to unsuitable soils or to lacking outlets. In the empirical application presented in section 4, such features are unlikely to occur. Our sample covers a limited geographical area and we only consider crops which can be profitably produced in this area.

$$\Pi_{it}(r) = \max_{\mathbf{s}} \left\{ \mathbf{s}'\boldsymbol{\pi}_{it} - C_{it}(\mathbf{s}) \text{ s.t. } \mathbf{s} \geq \mathbf{0}, \mathbf{s}'\boldsymbol{\iota} = 1 \text{ and } s_k = 0 \text{ if } k \in \mathcal{K}^0(r) \right\}. \quad (2b)$$

for $r \in \mathcal{R}$.

Second, the optimal production regime r_{it} is determined by comparing the regime specific expected profit levels while accounting for the production regime fixed costs. Accordingly, the expected profit maximizing production regime r_{it} is defined as the solution in r to a simple discrete maximization problem with:

$$r_{it} = \arg\max_{r \in \mathcal{R}} \left\{ \Pi_{it}(r) - D_{it}\left(\rho(\mathbf{s}_{it}(r)) \right) \right\} . \quad (3)$$

Assuming that optimal regime r_{it} is unique, optimal acreage choice $\mathbf{s}_{it}$ is obtained by combining equations (3) and (2a), with:

$$\mathbf{s}_{it} = \mathbf{s}_{it}(r_{it}) . \quad (4a)$$

Similarly, equations (3) and (2b) yield the expected profit level Π_{it}, with:

$$\Pi_{it} = \Pi_{it}(r_{it}) . \quad (4b)$$

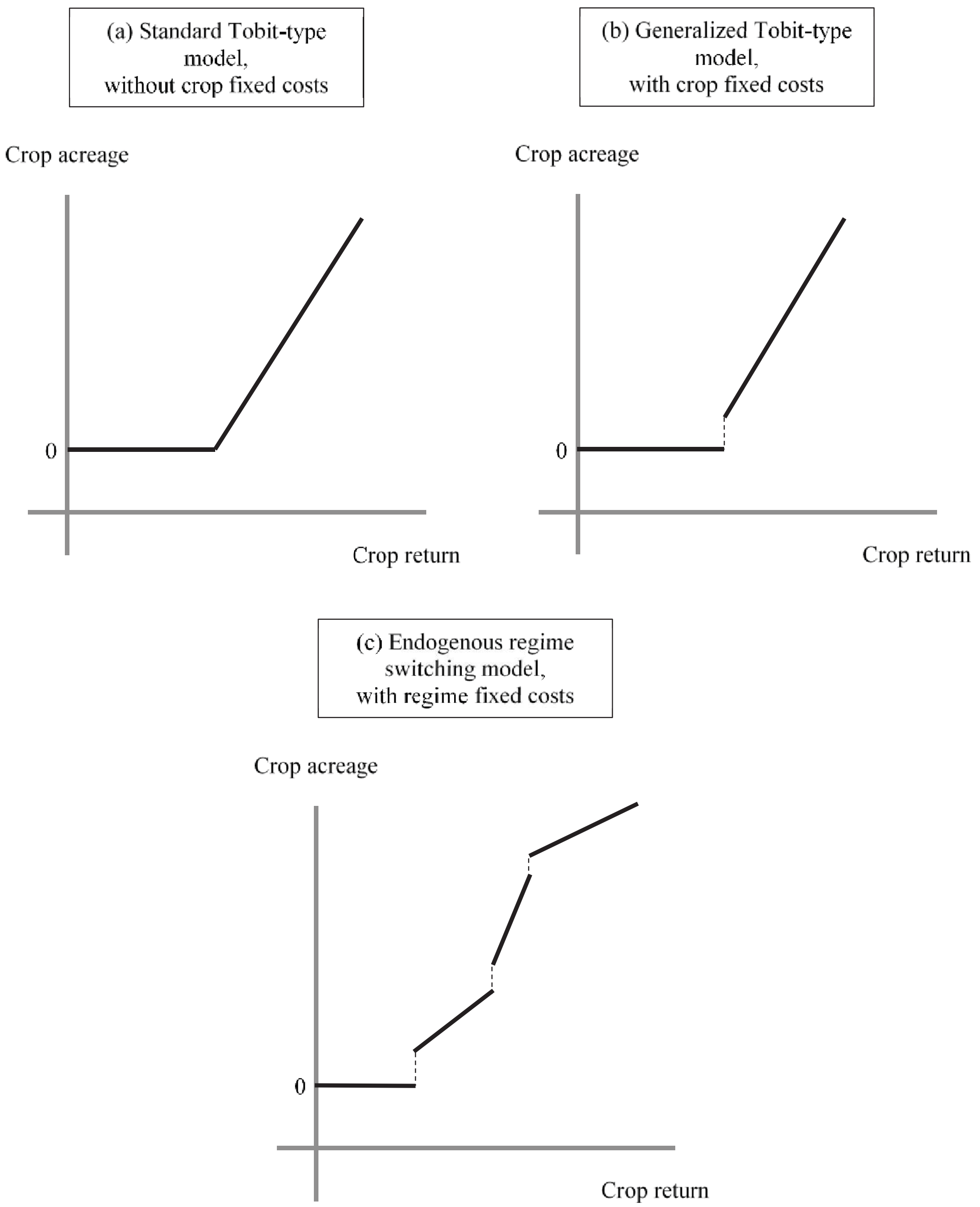
Regime specific acreage choices $\mathbf{s}_{it}(r)$ are derived from optimization problems that differ from one regime to the other due to nullity constraints on crop acreages. These constraints significantly impact how the acreage choices of the produced crops respond to market conditions. For instance, the regime r acreage choice, $\mathbf{s}_{it}(r)$, doesn't respond to changes in the expected returns of the crops not produced in regime r. Similarly, acreages of produced crops are expected to be more responsive to economic incentives in regimes containing numerous crops than in regimes containing only a few crops, crop acreage substitution opportunities being more limited with small crop sets.

2.2. Corners, kinks and jumps in acreage choice models

Our micro-economic crop acreage choice model is an example of ERS multivariate model with multiple regimes. To our knowledge, ERS models for multiple choices have been mostly used for demand systems, either for consumption goods (*e.g.*, Wales and Woodland, 1983; Lee and Pitt, 1986; Kao *et al.*, 2001; Millimet and Tchernis, 2008) or for production factors (*e.g.*, Lee and Pitt, 1987; Arndt, 1999: Chakir and Thomas, 2003). Most of these studies rely on the dual modelling framework proposed by Lee and Pitt (1986).

The main differences between the approaches that can be considered for handling null acreages in MEMC models are illustrated schematically in Figure 1. Panels (a)-(c) depict how the crop acreage of a given crop depends on its expected return according to three multi-crop acreage models. These models differ on how they handle null acreage choices – based either on ERS models or on CR systems – and on whether they account for crop or regime production fixed costs or not. Indeed, Figure 1 shows that this comparison is all about “corners”, “kinks” and “jumps”.

Figure 1. Typical multi-crop acreage models handling null crop acreages



Source : Authors.

Models that account for null acreages and don't account for crop production fixed costs are defined as systems of standard Tobit models (*e.g.*, Moore and Negri, 1992; Moore *et al.*, 1994). They define null acreages as corner solutions at zero. Their crop acreage models display one kink at the crop return level at which the non-negativity constraint of the considered crop just bind, as illustrated in panel (a).

Panel (b) depicts patterns allowed by models that account for null acreages based on CR systems as well as for crop production fixed costs. These crop acreage choice models display one kink and, potentially, a jump at the crop return level where farmers are indifferent between planting the considered crop or not. Being based on extensions of generalized Tobit models, recent CR-MEMC models (*e.g.*, Sckokai and Moro, 2006, 2009; Lacroix and Thomas, 2011; Bateman and Fezzi, 2011; Platoni *et al.*, 2012) implicitly account for production regime costs.

Crop acreage choices patterns allowed in our ERS-MEMC model are depicted in panel (c). Due to the effects of the regime choices on acreage choices, crop acreages may display several kinks. A kink occurs wherever changes in the expected return of the considered crop induce a regime switch. The first kink occurs at the crop return level above which farmers decide to plant the considered crop while others occur at regime switch points concerning the decision to produce or not to produce other crops. Our ERS-MEMC may also induce jumps at regime switch points, these jumps being due to threshold effects induced by regime fixed costs. According to our knowledge, this is the first MEMC model allowing such crop choice patterns.

2.3. Crop choices and MNL acreage choice models

The regime fixed cost considered in the maximization problem (3) determining the optimal regime r_{it} is $D_{it}\left(\rho(\mathbf{s}_{it}(r))\right)$ rather than simply $D_{it}(r)$. In effect, the production regime of $\mathbf{s}_{it}(r)$ may not be regime *r*, depending on the functional form chosen for the cost function $C_{it}(\mathbf{s})$. The regime of $\mathbf{s}_{it}(r)$ is only guaranteed to be a regime "included" in regime *r* as elements of $\mathbf{s}_{it}(r)$ may be null due to binding non-negativity constraints. The production regime of $\mathbf{s}_{it}(r)$ is regime *r* if and only if $s_{k,it}(r)$ is an interior solution to problem (2a) for any $k \in \mathcal{K}^{+}(r)$.

For instance, if $C_{it}(\mathbf{s})$ is quadratic in **s** then $s_{k,it}(r)$ is null if $\pi_{k,it}$ is sufficiently low. Moreover, neither crop acreage $\mathbf{s}_{it}(r)$ nor expected profit $\Pi_{it}(r)$ are obtained in analytical closed form in the quadratic case, precisely because elements of $\mathbf{s}_{it}(r)$ may be corner solutions at 0.

By contrast, the Multinomial Logit (MNL) crop acreage share models proposed by Carpentier and Letort (2014) appear especially convenient in this context.[9] This modelling framework relies on a family of acreage management cost functions ensuring that optimal crop acreage shares $\mathbf{s}_{it}(r)$ and expected profit levels $\Pi_{it}(r)$ satisfy two important conditions for any regime *r*.

First, these terms are obtained in analytical closed forms. For instance, if the acreage management cost function is assumed to have the linear-entropic functional form $C_{it}(\mathbf{s}) = \sum_{k \in \mathcal{K}^+(r)} s_k \beta^s_{k,it} + (\alpha^s_i)^{-1} \sum_{k \in \mathcal{K}^+(r)} s_k \ln s_k$ with $\alpha_i > 0$ then the regime specific acreage share vectors $\mathbf{s}_{it}(r)$ are given by Standard MNL acreage share models:

$$s_{k,it}(r) = \frac{j_k(r) \exp\left(\alpha^s_i(\pi_{k,it} - \beta^s_{k,it})\right)}{\sum_{\ell \in \mathcal{K}} j_\ell(r) \exp\left(\alpha^s_i(\pi_{\ell,it} - \beta^s_{\ell,it})\right)} \text{ for } k \in \mathcal{K}. \tag{5}$$

where function $j_k(r)$ indicates whether crop *k* belongs to regime *r* or not; with $j_k(r) = 1$ if $k \in \mathcal{K}^+(r)$ and $j_k(r) = 0$ otherwise. Second, it is easily seen from equation (5) that, for Standard MNL acreage share models, if crop *k* belongs to regime *r* then the optimal acreage share of crop *k* in regime *r* is ensured to be strictly positive. More generally, considering Standard or Nested MNL crop acreage models ensures that the production regime of $\mathbf{s}_{it}(r)$ is regime *r*.

The fact that $s_{k,it}(r)$ cannot be null means that null crop acreages are handled in a specific way in the MNL modelling framework. Crop acreage non-negativity constraints never bind when deriving MNL acreage share models.[10] These constraints just imply that the optimal acreage shares of the least profitable crops (acreage management cost included) are very small when they are much less profitable than other crops of the considered crop set.[11] The acreage shares of the least profitable crops may only become null when farmers choose their production regime. Farmers exclude these crops from their production plans when they can get higher expected profit level without planting them. Incidentally, this feature of MNL acreage choice models prevents their use in CR-MEMC models.

---

[9] Of course, choosing functional forms for their being convenient is unwarranted. Yet, their estimation being particularly challenging, all specifications of ERS models with multiple regimes that were used in empirical studies exploit, to some extent, properties of specific functional forms (e.g., Wales and Woodland, 1983; Lee and Pitt, 1986, 1987; Arndt et al, 1999). Also, other properties of MNL acreage share models make them empirically relevant for modelling production choices of arable crop producers (Carpentier and Letort, 2014).

[10] This property comes from properties of the entropy terms that appear in the acreage cost management functions leading to MNL acreage share models (Carpentier and Letort, 2014). Term $-s_k \ln s_k$ tends to 0 as s_k decreases to 0 (we have $s_k \ln s_k = 0$ if $s_k = 0$ according to a standard extension by continuity result) while its derivative in s_k tends to infinity as s_k decreases to 0.

[11] It is easily seen, from equation (5), that $s_{k,it}(r)$ decreases to 0 as $\pi_{k,it}$ decreases to $-\infty$.

3. ERS-MEMC model with regime specific fixed costs: micro-economic structure

This section presents the structure of the ERS-MEMC model considered in the empirical application presented in the next section. This model is composed, on the one hand, of yield supply functions, variable input demand functions and acreage share choice models for each produced crop, and on the other hand, of a probabilistic production regime choice model. This MEMC model can be interpreted as an extension to an ERS framework with regime fixed costs of the model proposed by Carpentier and Letort (2014).

As in Koutchadé *et al.* (2018) we adopt a random parameter approach for accounting for farmers' and farms' unobserved heterogeneity. We assume that the parameters of farmers' production choices, including those driving farmers' responses to economic incentives, are farm specific. Accordingly, the main aim of the estimation procedure is to recover their distribution across the farmers' population represented by the considered sample.

The considered ERS-MEMC model is presented in three steps. First, we present the production choice models defined at the crop level, *i.e.* the yield supply and variable input demand models. Second, we present the per regime acreage share choice models. Finally, we describe the production regime choice model. This presentation is organized following the structure of the model: yield supply and variable input demand models are used for defining expected crop return models. These models are then used for defining crop acreage share models, which are themselves used for defining the production regime choice model.

3.1. Yield supply and variable input demand models

We assume that farmers produce crop k from a variable input aggregate under a quadratic technological constraint. *I.e.*, we assume that the yield of crop k obtained by farmer i in year t is given by:

$$y_{k,it} = \beta_{k,it}^{y} - 1/2 \times (\alpha_{k,i}^{x})^{-1} (\beta_{k,it}^{x} - x_{k,it})^{2} \tag{6}$$

where $x_{k,it}$ denotes the variable input use level. Parameter $\alpha_{k,i}^{x}$ is required to be (strictly) positive for the production function to be (strictly) concave in $x_{k,it}$. It determines the extent to which the yield supply and the input demand of crop k respond to the input and crop prices.

Terms $\beta_{k,it}^{y}$ and $\beta_{k,it}^{x}$ have direct interpretations in the considered yield function. Term $\beta_{k,it}^{y}$ is the yield level that can be potentially achieved by farmer i in year t while $\beta_{k,it}^{x}$ is the input quantity required to achieve this potential yield level. These parameters are decomposed as

$\beta^{y}_{k,it} = \beta^{y}_{k,i} + (\boldsymbol{\delta}^{y}_{k,0})'\mathbf{c}^{y}_{k,it} + \varepsilon^{y}_{k,it}$ and $\beta^{x}_{k,it} = \beta^{x}_{k,i} + (\boldsymbol{\delta}^{x}_{k,0})'\mathbf{c}^{x}_{k,it} + \varepsilon^{x}_{k,it}$ where terms $\mathbf{c}^{y}_{k,it}$ and $\mathbf{c}^{x}_{k,it}$ are observed variable vectors used to control for observed farm heterogeneity (*i.e.*, farm size and capital endowment per unit of land) and climatic conditions (*i.e.*, temperature and rainfall). The $\beta^{y}_{k,it}$ and $\beta^{x}_{k,it}$ terms are farmer specific parameters aimed at capturing unobserved heterogeneity across farms and farmers. These terms, as well as the $\alpha^{x}_{k,i}$ random parameter, mainly capture three kinds of effects: those of the natural and material factor endowment of farms (*e.g.*, soil quality, machinery quality), of farmers' practice choices (*e.g.*, crop management practices, cropping systems) and of the skills of farmers. Terms $\varepsilon^{y}_{k,it}$ and $\varepsilon^{x}_{k,it}$ are standard error terms aimed to capture the effects on production of stochastic events (*e.g.*, climatic conditions, and pest and weed problems). We assume that farmer *i* is aware of the content of $\varepsilon^{x}_{k,it}$ when deciding his variable input uses.

Assuming that farmer *i* maximizes the expected return to variable input uses of each crop, we can easily derive the demand of the variable input for crop *k*:

$$y_{k,it} = \beta^{y}_{k,i} + (\boldsymbol{\delta}^{y}_{k,0})'\mathbf{c}^{y}_{k,it} - 1/2 \times \alpha^{x}_{k,i} w^{2}_{k,it} p^{-2}_{k,it} + \varepsilon^{y}_{k,it} \tag{7a}$$

and the corresponding yield supply:

$$x_{k,it} = \beta^{x}_{k,i} + (\boldsymbol{\delta}^{x}_{k,0})'\mathbf{c}^{x}_{k,it} - \alpha^{x}_{k,i} w_{k,it} p^{-1}_{k,it} + \varepsilon^{x}_{k,it} \ . \tag{7b}$$

Terms $p_{k,it}$ and $w_{k,it}$ respectively denote the expected output and input prices of crop *k*. Assuming that the expectations of $\varepsilon^{y}_{k,it}$ and $\varepsilon^{x}_{k,it}$ of farmer *i* are null at the beginning of the cropping season,[12] this farmer expects the following return to the variable input:

$$\pi_{k,it} = p_{k,it}\left(\beta^{y}_{k,i} + (\boldsymbol{\delta}^{y}_{k,0})'\mathbf{c}^{ys}_{k,it}\right) - w_{k,it}\left(\beta^{x}_{k,i} + (\boldsymbol{\delta}^{x}_{k,0})'\mathbf{c}^{xs}_{k,it}\right) + 1/2 \times \alpha^{x}_{k,i} w^{2}_{k,it} p^{-1}_{k,it} \tag{8}$$

for crop *k* when she/he chooses her/his acreage shares. Vector $(\mathbf{c}^{ys}_{k,it}, \mathbf{c}^{xs}_{k,it})$ is defined by replacing in vector $(\mathbf{c}^{y}_{k,it}, \mathbf{c}^{x}_{k,it})$ the climatic variables by their expectations.

3.2. Acreage share choice models

As discussed in Carpentier and Letort (2014), the Standard MNL crop acreage model given in equation (5) appears to be rather rigid because it treats the different crops symmetrically. Indeed, arable crops can often be grouped according to their competing for the use of quasi-

[12] As discussed below, this assumption can be relaxed, *e.g.* for accounting for potential correlations between the $\varepsilon^{y}_{k,it}$ and $\varepsilon^{x}_{k,it}$ error terms on the one hand, and the $\varepsilon^{s}_{k,it}$ error terms on the other hand.

fixed factors or according to their agronomic characteristics. The ERS-MEMC model considered in our application presented in the next section is based on a 3 level Nested Multinomial Logit (NMNL) acreage share model.

For sake of simplification, we consider a 2 level NMNL acreage share model in this section.[13] Crop set $\mathcal{K}$ is partitioned into G mutually exclusive groups of crops. Term $\mathcal{G} = \{1,...,G\}$ defines the considered group set. Group $g \in \mathcal{G}$ defines the crop subset $\mathcal{K}(g)$. Crops belonging to a same group are assumed to share similar agronomic characteristics and to compete more for farmers' limiting quantities of quasi-fixed factors than they compete with crops of other groups. The corresponding acreage management cost function is given by:

$$C_{it}(\mathbf{s}) = \sum\nolimits_{k\in\mathcal{K}} s_k \beta^s_{k,it} + \sum\nolimits_{g=1}^{G} (\alpha^s_i)^{-1} s_{(g)} \ln s_{(g)} + \sum\nolimits_{g\in\mathcal{G}} s_{(g)} (\alpha^s_{(g),i})^{-1} \sum\nolimits_{m\in\mathcal{K}(g)} s_{m|(g)} \ln s_{m|(g)} \quad (9)$$

where $s_{(g)}$ denotes the acreage share of group *g* and $s_{m,(g)}$ that of crop *m* in group *g*. Terms α^s_i and $\alpha^s_{(g),i}$ are farm specific parameters determining the flexibility of farmers' acreage choices.[14] The larger they are, the more the acreage share choice respond to economic incentives (because the less management costs matter). Condition $\alpha^s_{(g),i} \geq \alpha^s_i > 0$ is sufficient for cost function $C_{it}(\mathbf{s})$ to be strictly convex in **s**.

The linear terms of the cost function $C_{it}(\mathbf{s})$ are decomposed as $\beta^s_{k,it} = \beta^s_{k,i} + (\boldsymbol{\delta}^s_{k,0})'\mathbf{c}^s_{k,it} + \varepsilon^s_{k,it}$ where $\mathbf{z}^s_{k,it}$ are explanatory variable vectors used to control for observed heterogeneous factors and climatic events. Farm specific parameters $\beta^s_{k,i}$ account for unobserved heterogeneity effects. Error terms $\varepsilon^s_{k,it}$ capture the effects of stochastic variations of the cost due to random events such as unobserved interactions of climatic events and soil characteristics impacting the soil state at planting. Farmers are assumed to know these terms when choosing their acreages. Error terms $\varepsilon^s_{k,it}$ are assumed to be independent from the error terms of the yield supply and input demand equations, $\varepsilon^y_{k,it}$ and $\varepsilon^x_{k,it}$.

Farmers' optimal crop acreage choices as given by equation (2a) can be derived for any production regime. It suffices to solve the maximization problem given in equations (3). For instance, eight acreage share subsystems are considered in our empirical application, one for each production regime present in the data. Of course, the functional form of the derived acreage choice function depends on the subset of crops produced in the considered regime. Assuming that crop *k* belongs to group *g*, we obtain:

---

[13] The model used in our application is presented in the Online Appendix.

[14] We have we have $\alpha^s_{(g),i} = \alpha^s_i$ if group *g* contains a single crop.

$$s_{k,it}(r) = \frac{j_k(r)\exp\left(\alpha^s_{(g),i}(\pi_{k,it} - \beta^s_{k,it})\right)\left(\sum_{\ell\in\mathcal{K}(g)} j_\ell(r)\exp\left(\alpha^s_{(g),i}(\pi_{\ell,it} - \beta^s_{\ell,it})\right)\right)^{\alpha^s_i(\alpha^s_{(g),i})^{-1}-1}}{\sum_{h\in\mathcal{G}}\left(\sum_{\ell\in\mathcal{K}(h)} j_\ell(r)\exp\left(\alpha^s_{(h),i}(\pi_{\ell,it} - \beta^s_{\ell,it})\right)\right)^{\alpha^s_i(\alpha^s_{(h),i})^{-1}}} \tag{10}$$

and:

$$\Pi_{it}(r) = (\alpha^s_i)^{-1}\ln\sum_{h\in\mathcal{G}}\left(\sum_{\ell\in\mathcal{K}(h)} j_\ell(r)\exp\left(\alpha^s_{(h),i}(\pi_{\ell,it} - \beta^s_{\ell,it})\right)\right)^{\alpha^s_i(\alpha^s_{(h),i})^{-1}} . \tag{11}$$

Parameter α^s_i drives the land allocation to crop group acreages while parameters $\alpha^s_{(g),i}$ drive the allocation of the crop group acreages to crop acreages.

3.3. Production regime choice model

Observing that the regime specific optimal acreage choice $\mathbf{s}_{it}(r)$ necessarily belongs to regime *r* in the MNL case considered here, the regime specific expected profit levels $\Pi_{it}(r)$ can be used for defining a regime choice model based to the choice problem described in equation (4). Let define the regime fixed costs as $D_{it}(r) = d_i(r) - \sigma_i^{-1}e_{r,it}$. The farm specific parameters $d_i(r)$ aim to capture the effects of unobserved factors affecting the regime fixed costs. The error terms $e_{r,it}$ aim to capture the effects of stochastic factors and define the regime choice model as a probabilistic discrete choice model, with:

$$r_{it} = \arg\max_{r\in\mathcal{R}}\left\{\Pi_{it}(r) - d_i(r) + \sigma_i^{-1}e_{r,it}\right\}. \tag{12}$$

Scale parameter σ_i determines the extent to which the regime expected profit levels (*i.e.* the $\Pi_{it}(r) - d_i(r)$ terms) explain the production regime choice as regards to the effects of the $e_{r,it}$ idiosyncratic terms. The higher σ_i, the more the expected profit levels impact the observed regime choices.

Regime fixed costs $d_i(r)$ can be specified in different ways. These costs are expected to increase with the number of crops. Transaction costs and labor requirements related to a production regime increase with the number of crops produced in that regime. Indeed, one way to specify $d_i(r)$ is to consider a sum of fixed costs associated to each crop produced in the considered production regime, with $d_i(r) = \sum_{k\in\mathcal{K}^+(r)}\beta^c_{k,i}$ where $\beta^c_{k,i}$ is the fixed costs related to crop *k*. Interestingly, this specification allows computing the fixed costs of regimes which are not observed in the data. This is of particular interest for simulation purposes. For example, changes in market conditions can lead farms to adopt new production regimes. This regime fixed cost specification is used in our empirical application.

This specification of the regime fixed costs can be usefully compared with more general ones.

Farmers may purchase inputs specific to different crops from the same supplier, implying savings in the related transaction costs. Moreover, different crops may generate work peak loads during the same periods, implying that can concentrate their workload (or that of their employees) during these periods if they wish so. In these cases, the regime fixed costs are sub-additive in the crop fixed costs. One way to deal with this pattern consists of directly specifying these fixed costs as famers specific constant terms on a regime per regime basis, with $d_i(r) = d_{r,i}$ (given that the fixed cost of a "benchmark regime" needs to be normalized). Of course, the costs corresponding to regimes that are not observed in the data can't be recovered, thereby constraining the regime set that can be simulated to be equal to the one that is observed in the data.

3.4. Overall structure of the ERS-MEMC model

The ERS-MEMC model is composed of three main parts: a subsystem of yield supply and input demand equations (7), a set of per regime subsystems of acreage share equations (10) and a probabilistic regime choice model (12).

The set of dependent variables of this model contains the crop level production choices. These consist of the yield levels, input use levels and acreage shares of each crop that are produced by for farmer *i* in year *t*. These are collected in vector $(\mathbf{y}_{it}^{+}, \mathbf{x}_{it}^{+}, \mathbf{s}_{it}^{+})$. Production regime r_{it} is the last dependent variable of the model.

The set of explanatory variables contains crop prices, variable input prices and the control variable vectors used in the crop yield supply, input demand and acreage share equations for all crops. These variable are collected in vector $\mathbf{z}_{it}$, which defines the information set of the ERS-MEMC model.

The sole fixed parameters appearing in the model equations are the coefficients of the control variable coefficient vectors for all crops, which are collected in vector $\boldsymbol{\delta}_0$.

The considered ERS-MEMC model contains two main subsets of random components: a vector of random parameters and a vector of error terms.

Vector $\boldsymbol{\gamma}_i$ collects the farm specific parameters of the model, with $\boldsymbol{\gamma}_i = (\boldsymbol{\beta}_i, \boldsymbol{\alpha}_i, \sigma_i)$. This vector contains the potential yield parameters, the input requirement parameters, the cost function linear parameters and the crop fixed costs parameters for all crops. These random parameters are collected in vector $\boldsymbol{\beta}_i$. It also contains the input use flexibility parameters for all crops and the acreage choice flexibility parameters, which are collected in vector $\boldsymbol{\alpha}_i$. Finally, $\boldsymbol{\gamma}_i$ contains the scale parameter, σ_i, of the regime choice model.

In the error term vector $\boldsymbol{\varepsilon}_{it} = (\boldsymbol{\varepsilon}_{it}^{yx}, \boldsymbol{\varepsilon}_{it}^{s})$, sub-vector $\boldsymbol{\varepsilon}_{it}^{yx}$ collects the error terms error terms of crop yield supply and input demand equations for all crops while sub-vector $\boldsymbol{\varepsilon}_{it}^{s}$ collects those of the acreage share equations. Finally, vector $\mathbf{e}_{it}$ collects the error terms of the regime choice model (*i.e.*, $e_{r,it}$ for $r \in \mathcal{R}$).

4. ERS-MEMC model with regime specific fixed costs: estimation strategy

This section presents the main features of the estimation strategy adopted for estimating the ERS-MEMC model described above. As this model involve multiple endogenous regimes, considers numerous interrelated production choices and features random parameters, we impose parametric distributional assumptions on its random components (*i.e.* error terms and random parameters) that ensure its empirical tractability. We also impose simplifying assumptions regarding the dynamics of farmers' choices and the multi-crop production technology. These assumptions are presented and discussed first. Then, we present how the main parameters of interest of our ERS-MEMC model are recovered from the data. Finally, we briefly describe our estimation strategy. More specifically, we present the main estimation issues that we face when estimating our random parameter ERS-MEMC model and the approaches chosen for overcoming these issues. A detailed description of our estimation procedure is provided in a dedicated Online Appendix. This procedure combines techniques found in the micro-econometrics and computational statistics literatures.

4.1. Main probabilistic assumptions

We assume that terms $(\boldsymbol{\varepsilon}_{is}, \mathbf{e}_{is})$, $\boldsymbol{\gamma}_i$ and $\mathbf{z}_{it}$ are independently distributed for any pair (t,s). This implies that the explanatory variables vector, $\mathbf{z}_{it}$, is assumed to be (*i*) strictly exogenous with respect to the error term vectors and (*ii*) independent of the random parameters $\boldsymbol{\gamma}_i$. This latter assumption, which is standard in random parameter models, defines $\boldsymbol{\gamma}_i$ as a term that captures heterogeneity effects not captured by control variables $\mathbf{z}_{it}$.

We further assume that error term $(\boldsymbol{\varepsilon}_{it}, \mathbf{e}_{it})$ vectors are independently distributed across time. Combined with the fact that vector $\mathbf{z}_{it}$ doesn't contain any lagged endogenous variable, this serial independence assumption implies that our MEMC model can be interpreted as a reduced form model as regards the dynamic features of the modelled choices. Indeed, we hypothesize that random parameters $\boldsymbol{\gamma}_i$ capture the effects on farmers' production choices and performances of the stable crop rotation schemes that these farmers rely on.[15] Koutchadé *et al.* (2018) provide empirical results confirming this hypothesis with a sample of arable crop producers located in an area contiguous to the one considered in our application.

Finally, we assume that the error term vectors $\boldsymbol{\varepsilon}_{it}^{yx}$, $\boldsymbol{\varepsilon}_{it}^{s}$ and $\mathbf{e}_{it}$ are independent. Relaxing this independence assumption for $\boldsymbol{\varepsilon}_{it}^{s}$ and $\boldsymbol{\varepsilon}_{it}^{yx}$ is possible but significantly increases the estimation burden and Koutchadé *et al.* (2018), in a similar context, found that error terms $\boldsymbol{\varepsilon}_{it}^{s}$ and $\boldsymbol{\varepsilon}_{it}^{yx}$ were not significantly correlated.

4.2. Distributional assumptions

Random parameter vectors $\boldsymbol{\gamma}_i$ are assumed independent across farms. For sake of simplification, we assume here that these random parameter vectors are normally distributed, with $\boldsymbol{\gamma}_i \sim \mathcal{N}(\boldsymbol{\mu}_0, \boldsymbol{\Omega}_0)$. Various transformations of elements of $\boldsymbol{\gamma}_i$ actually allow for other distribution choices for these elements while keeping the multivariate structure of the probability distribution of $\boldsymbol{\gamma}_i$ (*e.g.*, Stanfield *et al.*, 1996). For example, considering log-transformations of $\boldsymbol{\alpha}_i$ and σ_i in $\boldsymbol{\gamma}_i$ implies that these random parameters, which are required to be positive, are jointly log-normality distributed. We used this log-transformation in the ESR-MEMC model used for our empirical application. Robustness checks demonstrated that other probability distribution choices have a limited impact on the main results.[16]

We make the usual assumptions stating that error term vectors $\boldsymbol{\varepsilon}_{it}$ are independent across farms (and years) and normally distributed, with $\boldsymbol{\varepsilon}_{it} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Psi}_0)$.[17]

Finally, we assume that the regime choice model error terms $e_{r,it}$ are independent across regimes and distributed according to a type I extreme value distribution. This assumption implies that the considered regime choice model is a standard Multinomial Logit discrete choice

[15] In that, we rely on well-known features of heterogeneous dynamic processes: those implying that empirically disentangling the effects of unobserved heterogeneity from those of unobserved persistent dynamic features is notably difficult. Accounting for dynamic features of multi-crop production technologies and of farmers' choices is challenging, and largely beyond the scope of this article.

[16] We tested specifications assuming that $\boldsymbol{\beta}_i$ is log-normally distributed and/or that $\boldsymbol{\alpha}_i$ follows a bounded Johnson distribution (*e.g.*, Stanfield *et al*, 1996).

[17] Matrix $\boldsymbol{\Psi}_0$ is block-diagonal under the assumption stating that $\boldsymbol{\varepsilon}_{it}^{s}$ and $\boldsymbol{\varepsilon}_{it}^{yx}$ are independent.

model conditionally on the scale parameter and on the regime specific expected profit levels and fixed costs. The corresponding conditional probability of the observed the regime choices is given by:

$$P(r_{it} \mid \boldsymbol{\varepsilon}_{it}^{s}, \mathbf{z}_{it}, \boldsymbol{\gamma}_i; \boldsymbol{\delta}_0) = \frac{\exp\left(\sigma_i(\Pi_{it}(r_{it}) - d_i(r_{it}))\right)}{\sum_{r\in\mathcal{R}} \exp\left(\sigma_i(\Pi_{it}(r) - d_i(r))\right)}. \tag{13}$$

This probability is defined as a function of $(\boldsymbol{\varepsilon}_{it}^{s}, \mathbf{z}_{it}, \boldsymbol{\gamma}_i; \boldsymbol{\delta}_0)$ because the vector of regime specific expected profit levels, $\Pi_{it}(r) - d_i(r)$ for $r \in \mathcal{R}$, is a function of all the terms contained in $(\boldsymbol{\varepsilon}_{it}^{s}, \mathbf{z}_{it}, \boldsymbol{\gamma}_i; \boldsymbol{\delta}_0)$, scale parameter σ_i excepted.

4.3. Identification

We consider here identification of the probability distribution of main random parameters of interest: the production choice flexibility parameters and the parameters of the regime choice model.

Under the considered assumptions the probability distribution of farmers' responses to economic incentives, $\boldsymbol{\alpha}_i$, are identified through two main channels. Identification of the probability distribution of the variable input use flexibility parameters, $\alpha_{k,i}^{x}$ for $k \in \mathcal{K}$, mostly relies on the variations of the corresponding input to crop price ratios. Identification of the probability distribution of the acreage choice flexibility parameters, α_i^{s} and $\alpha_{(g),i}^{s}$ for $g \in \mathcal{G}$, mainly relies on the variations of the expected crop return terms, $\pi_{k,it}$ for $k \in \mathcal{K}$. Importantly, the expected crop returns are defined as functions of random parameters (*i.e.*, $\beta_{k,i}^{y}$, $\beta_{k,i}^{x}$ and $\alpha_{k,i}^{x}$ for $k \in \mathcal{K}$) that may be correlated with the acreage choice flexibility parameters. The "full" variance-covariance matrix of the joint probability distribution of the random parameters $\boldsymbol{\gamma}_i$ takes into account these potential correlations.

Scale parameter σ_i, which is the random coefficient associated to the regime specific expected profit levels $\Pi_{it}(r)$ in the regime choice model, is mainly identified by the variations in these variables. Crop fixed costs $\beta_{k,i}^{c}$ are entailed in the regime fixed costs $d_i(r) = \sum_{k\in\mathcal{K}^{+}(r)} \beta_{k,i}^{c}$. Importantly, the fixed costs of the crops that are always produced cannot be identified because these crops are part of any regime present in the data. Therefore, the fixed costs of these crops are normalized at zero. The joint probability distribution of the identifiable crop fixed cost vector is mainly identified by the variations in the differences in the regime specific expected profit levels $\Pi_{it}(r)$ across the production regimes. The potential correlations between, on the one hand, the random parameters that are part of the expected profit levels and, on the other hand, the crop fixed costs and the scale parameter are taken into account in the distribution of $\boldsymbol{\gamma}_i$.

4.4. Estimation issues and sketch of the estimation procedure

The considered ERS-MCEM model being fully parametric, we consider a Maximum Likelihood (ML) estimator for efficiently estimating its parameters. These parameters are collected in $\boldsymbol{\theta}_0 = (\boldsymbol{\delta}_0, \boldsymbol{\Psi}_0, \boldsymbol{\mu}_0, \boldsymbol{\Omega}_0)$. Contribution of farmer *i* to the likelihood function of the model corresponds to the probability density function (pdf) of her/his sequence of production choices conditional on the sequence of exogenous variables characterizing this choice sequence. Assuming that the considered pdf is parameterized by $\boldsymbol{\eta}$, let function $f(\mathbf{u}\,|\,\mathbf{v};\boldsymbol{\eta})$ generically denotes the pdf of $\mathbf{u}_{it}$ conditional on $\mathbf{v}_{it} = \mathbf{v}$ at $\mathbf{u}_{it} = \mathbf{u}$. And let function $\varphi(\mathbf{u};\boldsymbol{\Omega})$ denote the pdf of $\mathcal{N}(\mathbf{0},\boldsymbol{\Omega})$ at $\mathbf{u}$. Given the probabilistic assumptions defining the parametric version of the random parameter ERS-MEMC model, contribution of farmer *i* to the likelihood function at $\boldsymbol{\theta}$ is given by:

$$\ell_i(\boldsymbol{\theta}) = \int \left(\prod_{t=1}^{T} f(\mathbf{y}_{it}^{+}, \mathbf{x}_{it}^{+}, \mathbf{s}_{it}^{+}, r_{it} \,|\, \mathbf{z}_{it}, \boldsymbol{\gamma}; \boldsymbol{\delta}, \boldsymbol{\Psi}) \right) \varphi(\boldsymbol{\gamma} - \boldsymbol{\mu}; \boldsymbol{\Omega}) d\boldsymbol{\gamma} . \tag{14}$$

Likelihood function $\ell_i(\boldsymbol{\theta})$ can be obtained neither analytically nor numerically due to its integration over the probability distribution of the random parameters $\boldsymbol{\gamma}_i$.

Micro-econometricians generally solve this problem by integrating $\ell_i(\boldsymbol{\theta})$ *via* direct simulation methods for computing Simulated ML (SML) estimators of $\boldsymbol{\theta}_0$. Yet, implementing this approach is particularly challenging with ERS-MEMC models due to the dimension of parameter $\boldsymbol{\theta}_0$ and the complexity of the simulated version of the likelihood functions $\ell_i(\boldsymbol{\theta})$. For instance, the ERS-MEMC model of our empirical application considers 22 production choices. It features 80 control variables, 37 random parameters and 20 error terms. Vector $\boldsymbol{\theta}_0$ contains 786 parameters while our dataset describes 40,192 observed production choices (16.5 per observation on average).

Integration of $\ell_i(\boldsymbol{\theta})$ over the random parameter distribution is thus the first estimation issue that we have to deal with. We compute the ML estimator of $\boldsymbol{\theta}_0$ by devising a Stochastic Approximate Expectation-Maximization (SAEM) algorithm. SAEM algorithms were proposed by Delyon *et al.* (1999) for computing ML estimators of models featuring continuous random parameters. These algorithms rely on simulation methods for integrating proxies of the sample log-likelihood of the considered model. They appear to use simulations more efficiently than competing alternatives (*e.g.*, McLachlan and Krishnan, 2007; Lavielle, 2014), which is a particularly relevant property when considering large samples, large multivariate models and/or large random parameter vectors. The structure of the SAEM algorithm that we propose for estimating random parameter ERS-MEMC models is described in the Appendix. Here, we consider its main step, the Maximization (M) step. This allows us to demonstrate the main advantages of our approach and, in the sequel, to illustrate the other two main estimation issues that we face.

At each of its iteration, the considered SAEM algorithm solves two maximization problems for updating estimates of $\boldsymbol{\theta}_0$. These problems have the form of weighted ML problems that are much simpler to solve than the corresponding SML problem. The first problem to be solved in the M step of our SAEM algorithm aims at updating the estimated value of $(\boldsymbol{\mu}_0, \boldsymbol{\Omega}_0)$, the parameter of the pdf of the model random parameters. It is of the form:

$$\max_{(\boldsymbol{\mu},\boldsymbol{\Omega})} \sum_{i=1}^{N} \sum_{j=1}^{J} \tilde{\eta}_i^j \ln \varphi(\tilde{\boldsymbol{\gamma}}_i^i - \boldsymbol{\mu}; \boldsymbol{\Omega}) \tag{15}$$

where terms $\tilde{\boldsymbol{\gamma}}_i^j$ are random draws of $\boldsymbol{\gamma}_i$ from a pdf defined by the preceding iteration results and $\tilde{\eta}_i^j$ are weighting terms attached to these draws. The solution in $\boldsymbol{\mu}$ is the empirical weighted mean of the random draws $\tilde{\boldsymbol{\gamma}}_i^j$ while the solution in $\boldsymbol{\Omega}$ is their empirical weighted variance-covariance matrix.

The second part of the M step of our SAEM algorithm updates the estimate of $(\boldsymbol{\delta}_0, \boldsymbol{\Psi}_0)$. It considers functions of the form $\tilde{W}(\boldsymbol{\delta},\boldsymbol{\Psi}) = \sum_{t=1}^{T} \sum_{i=1}^{N} \sum_{j=1}^{J} \tilde{\eta}_i^j \ln f(\mathbf{w}_{it}^+, r_{it} \mid \mathbf{z}_{it}, \tilde{\boldsymbol{\gamma}}_i^j; \boldsymbol{\delta}, \boldsymbol{\Psi})$ where $\mathbf{w}_{it}^+ = (\mathbf{y}_{it}^+, \mathbf{x}_{it}^+, \mathbf{s}_{it}^+)$. These functions have the functional forms of a weighted log-likelihood function of the production choices of a sample of "simulated farmers". Assuming that $(\hat{\boldsymbol{\delta}}, \hat{\boldsymbol{\Psi}})$ is the preceding iteration estimate of $(\boldsymbol{\delta}_0, \boldsymbol{\Psi}_0)$, it consists of solving either of the two following problems (*a*) $\max_{(\boldsymbol{\delta},\boldsymbol{\Psi})} \tilde{W}(\boldsymbol{\delta},\boldsymbol{\Psi})$ or (*b*) find $(\boldsymbol{\delta},\boldsymbol{\Psi})$ such that $\tilde{W}(\boldsymbol{\delta},\boldsymbol{\Psi}) \geq \tilde{W}(\hat{\boldsymbol{\delta}},\hat{\boldsymbol{\Psi}})$. Unfortunately, solving problem (*a*) or even simpler search problem (*b*) is difficult due to the complexity of the conditional likelihood function $f(\mathbf{w}_{it}^+, r_{it} \mid \mathbf{z}_{it}, \boldsymbol{\gamma}; \boldsymbol{\delta}, \boldsymbol{\Psi})$.

Decomposing this function demonstrate that the problem is indeed twofold. Using Bayes's law and the structure and distributional assumptions of the ERS-MCEM model, we obtain:

$$f(\mathbf{w}_{it}^+, r_{it} \mid \mathbf{z}_{it}, \boldsymbol{\gamma}; \boldsymbol{\delta}, \boldsymbol{\Psi}) = P(r_{it} \mid \mathbf{s}_{it}^+, \mathbf{z}_{it}, \boldsymbol{\gamma}; \boldsymbol{\delta}, \boldsymbol{\Psi}) f(\mathbf{w}_{it}^+ \mid \mathbf{z}_{it}, \boldsymbol{\gamma}; \boldsymbol{\delta}, \boldsymbol{\Psi}). \tag{16}$$

Vector $\mathbf{s}_{it}^{+}$ collects the acreage shares of the crops produced in regime r_{it}.[18] Function $f(\mathbf{w}_{it}^{+} | \mathbf{z}_{it}, \boldsymbol{\gamma}; \boldsymbol{\delta}, \boldsymbol{\Psi})$ is the likelihood of crop level choice vector $\mathbf{w}_{it}^{+}$ conditional on $(\mathbf{z}_{it}, \boldsymbol{\gamma}_i = \boldsymbol{\gamma})$ and $P(r_{it} | \mathbf{s}_{it}^{+}, \mathbf{z}_{it}, \boldsymbol{\gamma}; \boldsymbol{\delta}, \boldsymbol{\Psi})$ is the probability function of regime r_{it} conditional on $(\mathbf{s}_{it}^{+}, \mathbf{z}_{it}, \boldsymbol{\gamma}_i = \boldsymbol{\gamma})$. Yet, both functions raise estimation issues.

Given the structure of our MEMC model, function $\ln f(\mathbf{w}_{it}^{+} | \mathbf{z}_{it}, \boldsymbol{\gamma}; \boldsymbol{\delta}, \boldsymbol{\Psi})$ is the likelihood function of a Gaussian Seemingly Unrelated Regression (SUR) system with observations missing at random (up an additive term that doesn't depend on $(\boldsymbol{\delta}, \boldsymbol{\Psi})$). The missing observations are the yield level, input use and acreage share of the crops that are not produced in regime r_{it}. Ruud (1991) discussed the use of Expectation-Maximization (EM) algorithms for alleviating the computation burden of ML estimators of models based on latent Gaussian SUR systems with missing observations. Based on Ruud's insights we devised an EM type approach for updating the estimates of $(\boldsymbol{\delta}_0, \boldsymbol{\Psi}_0)$ in the M step of our SAEM algorithm.

Our last main estimation issue is due to the computation of the regime choice probability function $P(r_{it} | \mathbf{s}_{it}^{+}, \mathbf{z}_{it}, \boldsymbol{\gamma}; \boldsymbol{\delta}, \boldsymbol{\Psi})$. Given the structure of our MEMC model, this probability function can be defined as a function of the error terms of the acreage share equations. Let vector $\boldsymbol{\varepsilon}_{it}^{s,+}$ collect the error terms of the acreage share models of the crops produced in regime r_{it} and vector $\boldsymbol{\varepsilon}_{it}^{s,0}$ collect those of the crops that are not produced. Vector $\boldsymbol{\varepsilon}_{it}^{s,+}$ can be recovered from the acreage share model and the data, the observed crop acreages of the produced crops in particular. Let function $\hat{\boldsymbol{\varepsilon}}_{it}^{s,+}(\boldsymbol{\gamma}, \boldsymbol{\delta})$ denote the residual function corresponding to error term $\boldsymbol{\varepsilon}_{it}^{s,+}$ The structure of our MEMC model and equation (13) yield:

$$P(r_{it} | \mathbf{s}_{it}^{+}, \mathbf{z}_{it}, \boldsymbol{\gamma}; \boldsymbol{\delta}, \boldsymbol{\Psi}) = \int P(r_{it} | \boldsymbol{\gamma}, \mathbf{z}_{it}, \boldsymbol{\varepsilon}_{it}^{s}; \boldsymbol{\delta}) f\left(\boldsymbol{\varepsilon}_{it}^{s,0} | \hat{\boldsymbol{\varepsilon}}_{it}^{s,+}(\boldsymbol{\gamma}, \boldsymbol{\delta}); \boldsymbol{\delta}, \boldsymbol{\Psi}\right) d\boldsymbol{\varepsilon}_{it}^{s,0} \qquad (17)$$

where $f(\boldsymbol{\varepsilon}_{it}^{s,0} | \hat{\boldsymbol{\varepsilon}}^{s,+}; \boldsymbol{\delta}_0, \boldsymbol{\Psi}_0)$ denotes the pdf of $\boldsymbol{\varepsilon}_{it}^{s,0}$ conditional on $\boldsymbol{\varepsilon}_{it}^{s,+} = \hat{\boldsymbol{\varepsilon}}^{s,+}$, which is normal. Vector $\boldsymbol{\varepsilon}_{it}^{s,0}$ must be considered as missing variables in the estimation process because it cannot be recovered by combining the model and the data. The Multinomial Logit functional form of function $P(r_{it} | \boldsymbol{\gamma}, \mathbf{z}_{it}, \boldsymbol{\varepsilon}_{it}^{s}; \boldsymbol{\delta})$ prevents its integration over the probability distribution of $\boldsymbol{\varepsilon}_{it}^{s,0}$, either analytically or numerically. Building on the work of Harding and Hausman (2007), we use Laplace approximates of the regime choice probability functions $P(r_{it} | \mathbf{s}_{it}^{+}, \mathbf{z}_{it}, \boldsymbol{\gamma}; \boldsymbol{\delta}, \boldsymbol{\Psi})$ for

[18] Yield supply and input demand levels, $(\mathbf{y}_{it}^{+}, \mathbf{x}_{it}^{+})$, and regime choices, r_{it}, are independent conditionally on acreage choices, exogenous variables and random parameters, $(\mathbf{s}_{it}^{+}, \mathbf{z}_{it}, \boldsymbol{\gamma}_i = \boldsymbol{\gamma})$, since error terms $(\boldsymbol{\varepsilon}_{it}^{y}, \boldsymbol{\varepsilon}_{it}^{x})$, $\boldsymbol{\varepsilon}_{it}^{s}$ and $\mathbf{e}_{it}$ are assumed to be mutually independent.

computing the likelihood function of our model.[19]

The fact that production regime choices and acreage choices depend on $\boldsymbol{\varepsilon}_{it}^{s}$ constitutes the first source of endogeneity of the regime choices in our ERS-MEMC model.[20] Random parameter $\boldsymbol{\gamma}_i$ constitutes a supplementary source of regime choice endogeneity in our ERS-MEMC model.

5. Empirical application: crop diversification of French arable crop producers

This section presents an application aimed to illustrate the empirical tractability of our modelling approach as well as to demonstrate the role of crop set choices in analyzes of farmers' production choices.

5.1. Data and model specification details

The model is estimated on an unbalanced panel data set containing 2276 observations of 415 French grain crop producers in the North and North-East of France, over the years 2006 to 2011. This sample has been extracted from data provided by an accounting agency located in the French territorial division *La Marne*. It contains detailed information about crop production for each farm (acreages, yields, input uses and crop prices at the farm gate). We consider seven crops: sugar beet, alfalfa, protein pea, rapeseed, winter wheat, corn and spring barley, which represent more than 80% of the total acreage in the considered area.[21]

The variable input aggregate accounts for the use of fertilizers, pesticides and seeds. The corresponding price index is computed as a standard Tornqvist index. When a farmer doesn't produce a crop the corresponding output and input prices are unobserved. These missing prices were approximated by the yearly average of the corresponding observed prices. All prices are deflated by the hired production services price index (base 1 in 2006) obtained from the French

[19] This approach relies on a second order Taylor expansion in $\boldsymbol{\varepsilon}_{it}^{s,0}$ of function $P(r_{it} \mid \boldsymbol{\gamma}, \mathbf{z}_{it}, \boldsymbol{\varepsilon}_{it}^{s}; \boldsymbol{\delta}) f\left(\boldsymbol{\varepsilon}_{it}^{s,0} \mid \hat{\boldsymbol{\varepsilon}}_{it}^{s,+}(\boldsymbol{\gamma}, \boldsymbol{\delta}); \boldsymbol{\delta}, \boldsymbol{\Psi}\right)$ around an optimally chosen value of $\boldsymbol{\varepsilon}_{it}^{s,0}$. Using simulation methods for integrating function $P(r_{it} \mid \mathbf{s}_{it}^{+}, \mathbf{z}_{it}, \boldsymbol{\gamma}; \boldsymbol{\delta}, \boldsymbol{\Psi})$ would be inconvenient in our case due to our using such methods for dealing with random parameters.

[20] Indeed, the endogeneity issues raised by $\boldsymbol{\varepsilon}_{it}^{s}$ in our ERS-MEMC model are analogous, from an econometric viewpoint, to those raised by the demand function error terms in demand systems with binding non-negativity constraints (*e.g.*, Wales and Woodland, 1983; Lee and Pitt; 1986).

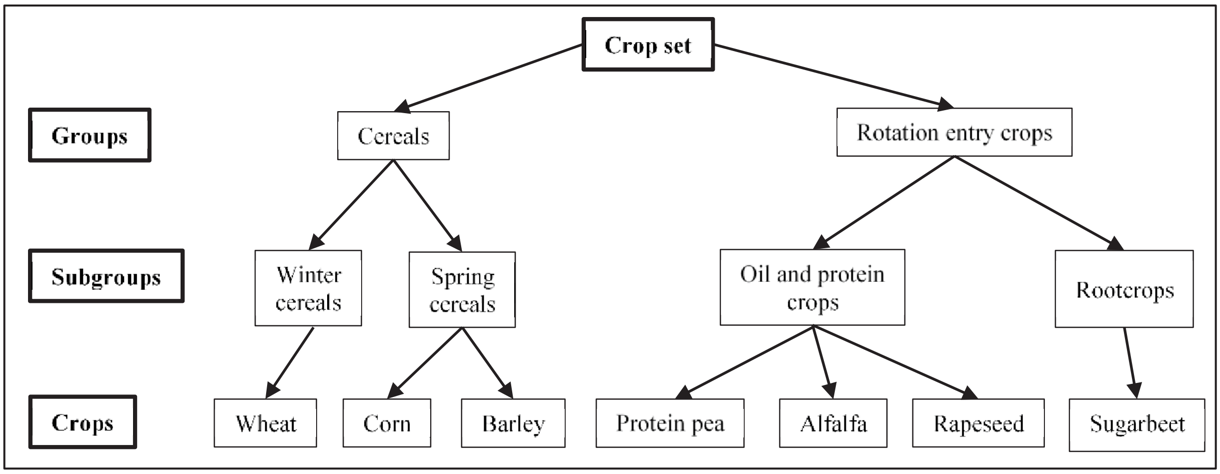
[21] The EU sugar beet subsidy scheme requires limited adjustments in our application because the actual sugar beet production largely exceeds the subsidized quota for all sugar beet producers of our sample.

department of Agriculture. This aggregated price index mainly depends on the price indices of machinery, fuel and hired labor, the main inputs involved in the implicit acreage management cost function. Climatic variables are provided at the municipality level by Météo France, the French national meteorological service.

Farmers' crop price expectations are defined by the corresponding lagged prices, according to a naïve anticipation scheme. Robustness checks demonstrate that anticipation scheme choices mostly impact estimates of the probability distribution of input use flexibility parameters $\alpha_{k,i}^{x}$, with very limited effects on our main results.

Figure 2 depicts the three levels nesting structure that we adopt for the seven crops. In a first level we distinguish a cereal group composed of wheat, corn and barley, and a group of rotation entry crops: sugar beet, alfalfa, peas and rapeseed. This structure is intended to reflect the basic rotation scheme of grain and industrial crop producers in France. In a second level, the cereal group is split into two subgroups: winter cereals on the one hand and spring cereals on the other hand, in order to account for the differences in planting seasons. The 'rotation entry crop' group is split into an 'oilseeds and protein crops' subgroup and a subgroup including only sugar beet (the only root crop considered here). Wheat, which is the only winter cereal, is used as the benchmark crop. Based on these seven crops, 127 regimes could theoretically be chosen by farmers. The 8 most frequently observed regimes, out of 78 regimes present in the original dataset, were considered for selecting our estimation sample.[22]

Figure 2. Nesting structure of the acreage choice model



Source : Authors.

[22] Considering a small regime set allowed us to estimate our ERS-MEMC model with regime specific fixed costs that are not defined as sums of crop fixed costs. This specification of the regime fixed costs is more flexible but only yields a modest improvement in the fit performance of the regime choice model.

All farmers grow winter cereals, (spring) barley, (winter) rapeseed and most of them (91.7%) grow at least two additional crops. The most frequent regimes in the sample (regimes 2, 3 and 4) include five or six crops. Table 1 provides descriptive statistics concerning the production regimes observed in the data. Most farmers adopt different production regimes over the 6 years of our sample: only 8 out of 415 farmers have not changed their production regime. The average gross margins associated to each regime are reported in the last column of Table 1. An interesting feature appears here: the most frequently chosen regimes are not the ones that lead to the highest average gross margin per hectare. For instance, regime 2 – which excludes corn – is characterized by the highest observed gross margin on average, but has been adopted in only 21.5% of the observations. This comes to illustrate the fact that farmers' choices of production regime are driven by factors other than gross returns, such as the acreage management and regime fixed costs represented in our model.

Because we assume that regime costs are equal to a sum of fixed costs associated to each crop produced in the considered regime, the fixed costs associated to winter cereals, spring barley and rapeseed, which are always produced in our sample, are set to zero for normalization purpose. Interestingly, our data configuration illustrates an important advantage of this regime fixed cost specification. According to Table 1, the less frequently produced crop (*i.e.*, corn) is produced in at least 24% of our observations while 3 production regimes (*i.e.*, regimes 5, 6 and 8) out of 8 are adopted in less than 3% of our observations. The probability distribution of fixed costs cannot be estimated accurately with our dataset on a pure per regime basis. But, that of crop fixed costs can be.

Table 1. Descriptive Statistics

Regime number	Average crops acreage shares per regime							Regime frequency	Average gross margin (€/ha)[b]
	Winter wheat	*Corn*	*Spring Barley*	*Sugar beet*	*Alfalfa*	*Protein pea*	*Rapeseed*		
1	0.38	0.07	0.15	0.12	0.09	0.06	0.13	6.6%	953
2	0.37		0.16	0.15	0.11	0.06	0.15	21.5%	1014
3	0.38	0.07	0.17	0.14	0.10		0.14	11.8%	930
4	0.37		0.20	0.16	0.11		0.16	48.6%	1007
5	0.41	0.14	0.19	0.10			0.15	2.8%	989
6	0.50	0.14	0.14				0.22	2.5%	825
7	0.44		0.23	0.14			0.19	4.9%	970
8	0.58		0.15				0.27	1.3%	738
Production frequency	100%	24%	100%	96%	88%	28%	100%		
Average acreage share[a]	0.38 (0.09)	0.02 (0.05)	0.18 (0.07)	0.15 (0.06)	0.10 (0.05)	0.02 (0.03)	0.16 (0.06)		
Average acreage share[a] if produced[a]	0.38 (0.09)	0.08 (0.07)	0.18 (0.07)	0.15 (0.06)	0.11 (0.04)	0.06 (0.03)	0.16 (0.06)		
Average gross margin (€/ha)[a,b]	843 (327)	872 (449)	756 (287)	1789 (379)	562 (286)	663 (269)	843 (311)		
Average yield (t/ha)[a]	8.58[b] (0.88)	9.23 (1.73)	6.82 (1.21)	95.19 (13.01)	12.62 (1.92)	4.72 (1.28)	3.89 (0.64)		
Average price (€/t)[a]	149[b] (31)	131 (34)	155 (35)	25 (3)	72 (15)	198 (25)	323 (64)		
Average fertilization and crop protection costs[a]	431 (91)	308 (74)	294 (70)	547 (126)	350 (125)	246 (66)	415 (83)		

Notes: [a] Empirical standard deviation in parentheses, [b] Off-quota price of sugar beet.

5.2. Estimation results

The parameter estimates of the yield, input demand, acreage shares and regime choice equations are reported in Tables 2 to 4. As shown in Table 2, the expectations of random parameters representing potential yields, $\beta_{k,i}^{y}$, are precisely estimated for all crops and their values lie in reasonable ranges regarding the average yields observed in the sample (Table 1). More importantly, the variances of their distributions are also statistically different from zero for all crops. These parameters thus significantly vary across farms, despite the fact that we control for observed factors characterizing farm heterogeneity (land and capital endowments and climatic conditions). This comes to illustrate the importance of unobserved farm heterogeneity in our sample.

Table 2. Selected Parameter Estimates of Yield Supply and Input Demand Models

	Winter wheat	*Corn*	*Spring barley*	*Sugar beet*	*Alfalfa*	*Protein pea*	*Rape-seed*
Yield supply model							
Error term $\varepsilon_{k,it}^{y}$							
Standard deviation	0.66*	1.83*	0.95*	9.70*	2.96*	1.72*	0.49*
	(0.02)	(0.07)	(0.02)	(0.02)	(0.02)	(0.03)	(0.016)
Potential yield $\beta_{k,i}^{y}$							
Mean	8.71*	9.06*	6.81*	95.60*	12.23*	4.15*	4.04*
	(0.02)	(0.04)	(0.02)	(0.32)	(0.04)	(0.03)	(0.01)
Standard deviation	0.26*	0.65*	0.33*	5.7*	0.69*	0.51*	0.24*
	(0.01)	(0.03)	(0.01)	(0.17)	(0.02)	(0.02)	(0.01)
Input demand model							
Error term $\varepsilon_{k,it}^{x}$							
Standard deviation	0.52*	0.59*	0.41*	0.84*	0.88*	0.60*	0.58*
	(0.01)	(0.02)	(0.01)	(0.02)	(0.02)	(0.02)	(0.01)
Input requirement $\beta_{k,i}^{x}$							
Mean	4.36*	2.57*	2.92*	5.44*	3.15*	2.29*	4.44*
	(0.02)	(0.02)	(0.01)	(0.03)	(0.03)	(0.02)	(0.02)
Standard deviation	0.37*	0.33*	0.24*	0.54*	0.44*	0.37*	0.41*
	(0.02)	(0.01)	(0.01)	(0.02)	(0.01)	(0.01)	(0.02)
Input use flexibility $\alpha_{k,i}^{x}$							
Mean	0.43*	0.08*	0.30*	0.49*	0.25*	0.33*	0.79*
	(0.01)	(0.00)	(0.00)	(0.03)	(0.00)	(0.01)	(0.02)
Standard deviation	0.13*	0.09*	0.05*	0.58*	0.02	0.18*	0.31*
	(0.01)	(0.04)	(0.00)	(0.06)	(0.03)	(0.01)	(0.01)

Notes: Estimated standard errors of the ML estimator are in parentheses. Asterisk (*) denotes a statistically significant parameter at the 5% level.

The parameter estimates of the input demand equations, also reported in Table 2, confirm this result: the probability distribution of their farm specific parameters is precisely estimated and displays significant heterogeneity. This is true for the random intercepts $\beta_{k,i}^{x}$ (the input use requirement) but also for the random slope parameters, $\alpha_{k,i}^{x}$, which represents the response of farmers to change in netput prices.

Turning to the parameter estimates of the acreage share equations in Table 3, again, the expectations and variance of random parameters are precisely estimated. Ranges of expectations of the acreage flexibility parameters are theoretically consistent. Conditions $\alpha_{m|(g),i}^{s} \geq \alpha_{(g),i}^{s} \geq \alpha_{i}^{s} > 0$ hold on average. These are sufficient conditions for the acreage model to be well-behaved.

Table 3. Selected Parameter Estimates of the Acreage Share Models

Crop level random terms	*Winter wheat*	*Corn*	*Spring barley*	*Sugar beet*	*Alfalfa*	*Protein pea*	*Rape--seed*
Error term $\varepsilon_{k,it}^{s}$							
Standard deviation	0	11.12* (0.38)	9.91* (0.19)	6.25* (0.13)	6.77* (0.15)	8.56* (0.28)	7.09* (0.16)
Acreage share shifters $\beta_{k,i}^{s}$							
Mean	0	17.41* (0.73)	13.88* (0.37)	24.51* (0.23)	11.15* (0.24)	18.78* (0.37)	11.07* (0.24)
Standard deviation	0	3.92* (0.02)	4.19* (0.02)	3.96* (0.03)	2.70* (0.06)	2.62* (0.01)	2.20* (0.01)

Acreage choice flexibility parameters	***Level 1*** α_{i}^{s}	***Level 2 (groups)*** $\alpha_{(g),i}^{s}$		***Level 3 (subgroups)*** $\alpha_{n\|(g),i}^{s}$	
		Cereals	**Rotation heads**	**Spring cereals**	**Oil and protein crops**
	Cereals *vs* rotation heads	Spring cereals *vs* winter cereals	Sugar beet *vs* oil and protein crops	Corn *vs* spring barley	Rapeseed *vs* protein pea *vs* alfalfa
Mean	0.046* (0.001)	0.053* (0.001)	0.073* (0.001)	0.530* (0.029)	0.11* (0.002)
Standard deviation	0.015* (0.001)	0.013* (0.001)	0.025* (0.001)	0.640* (0.029)	0.020* (0.002)

Notes: Estimated standard errors of the ML estimator are in parentheses. Asterisk (*) denotes a statistically significant parameter at the 5% level.

Finally, as shown in Table 4, the regime costs associated to crops, $d_{k,i}^{c}$, and the scale parameter, σ_i, of the regime choice equation are significantly estimated and heterogeneous across the sample. The mean value of the scale parameter, 1.80, is large, reflecting the importance of regime profit and fixed cost levels in production regime choices. Simulation results provided in the next sub-section illustrate this point. Estimated mean fixed cost of alfalfa is negative on average. Two main reasons might explain this result. First, alfalfa is planted for at least two years. This crop requires farmers' intervention mostly at planting and harvesting. In the *Marne* region, the alfalfa downstream (dehydration) industry generally takes on harvest operations, which comes to decrease farmers' workload significantly. Second, being a legume alfalfa exhibits good agronomic properties, especially when used as a previous crop for cereals. Crop fixed cost estimates should, however, be considered cautiously given their high variability across farms.

Once we have estimated the parameters characterizing the distribution of the random parameters $\boldsymbol{\gamma}_i$, we can "statistically calibrate" those parameters for each farmer in our sample and thus obtain a set of farmer specific "calibrated" models that can then be used for simulation purposes (Koutchadé *et al.*, 2018). In this study, the specific parameter $\boldsymbol{\gamma}_i$ of farm *i* is calibrated as the mode of its (simulated) probability distribution conditional on $(\mathbf{y}_{it}^{+}, \mathbf{x}_{it}^{+}, \mathbf{s}_{it}^{+}, r_{it}, \mathbf{z}_{it})$ for $t = 1,...,T$ (*i.e.* according to a ML 'calibration' criterion conditionally on what is known about farm *i* in the data). One interesting feature is that this procedure also allows us to calibrate the parameters of the yield, input demand and acreage equations corresponding to crops that have not been grown by the considered farmer as well as farmer specific regime fixed costs for regimes that have never been chosen by the considered farmer.

Table 4. Parameter Estimates of Regime Choice Models

	Crop fixed costs $\beta_{k,i}^{c}$							Scale parameter σ_i
	Winter wheat	***Corn***	***Spring barley***	***Sugar beet***	***Alfalfa***	***Peas***	***Rapeseed***	
Mean[a]	0	3.80* (0.24)	0	0.30* (0.12)	-4.70* (0.28)	1.30* (0.04)	0	1.80* (0.07)
Std dev[a]	0	4.16* (0.10)	0	2.22* (0.05)	4.40* (0.11)	0.67* (0.01)	0	1.40* (0.07)

Notes: Estimated standard deviation of the estimator in parentheses. Asterisk (*) denotes a statistically non null parameter at the 5% level.

Table 5. Fitting Criteria (*Sim-R²*)

	Winter wheat	*Corn*	*Spring barley*	*Sugar beet*	*Alfalfa*	*Peas*	*Rape- -seed*
Yield supply models	0.37	0.24	0.35	0.42	0.28	0.39	0.45
Input demand models	0.44	0.30	0.40	0.34	0.30	0.43	0.40
Acreage share models		0.57	0.34	0.83	0.70	0.53	0.41

The estimated farmer specific models allow us to compute fitting criteria, Sim-R^2, which are reported in Table 5. The Sim-R^2 criterion measures the quality of the prediction of the observed choices of farmers by the estimated models. Its construction is analogous to that of the R^2 criterion of the standard linear regression model: for a given choice variable and a given model, the Sim-R^2 criterion is defined as the ratio of the empirical variance of the prediction of this variable to the empirical variance of the observed variable.

These estimated criteria tend to show that the proposed model offers a satisfactory fit to our data.[23] Using the estimated farmer specific models to predict the regime choices observed in our data, we find our model to exhibit a relatively good predictive power with 72.4% of regime choices correctly predicted. Importantly, our investigations on this issue tend to demonstrate that our results are robust to various distributional assumptions related to the model random parameters.

5.3. Simulation results

The structure of the proposed ERS multi-crop micro-econometric model allows to investigate the relative importance of the main drivers of production regime choices. For that purpose, we consider the simulation model obtained from the estimated one by calibrating the farm specific parameters for each farm of our sample. Then we use this simulation model for investigating the prediction power of three elements of the regime choice models: the weighted sum of the expected crop gross returns $\boldsymbol{\pi}'_{it}\mathbf{s}_{it}(r)$, the acreage management costs $C_{it}\left(\mathbf{s}_{it}(r)\right)$ and the regime fixed costs d_i^r for $r \in \mathcal{R}$. We simulate the regime choices according to each of these elements as well as combinations of these elements, and then confront them, on average, with the observed regime choices. Taken together these simulation results confirm that regime fixed

[23] Much better fit levels are obtained for crop supply, acreage and input demand model defined at the farm level, mostly due to the explanatory power of the cropland area variable.

costs matter, but mainly in combination with the other drivers of the regime choice model. The maximization of gross margins, or the minimization of acreage management costs or regime fixed cost alone leads to predictions of regime choices that are strongly biased on average. Considering pairs of these choice criteria only slightly improve the predictions, while considering together these three criteria unsurprisingly provides predicted choices very close, on average, to the observed ones.

To illustrate the relevance of the approach we propose to deal with corner solutions in acreage choices, we simulate the impacts of changes in expected crop prices on acreage choices. As acreage price elasticities play a crucial role in this type of exercise, we present them first. In our ERS-MEMC model these elasticities account for the impact of crop prices on both acreages within any given regime and switch in production regimes. These two effects can be distinguished by generalizing, to a multiple regime case, the decomposition proposed by McDonald and Moffit (1980) for standard Tobit models. The average acreage own price elasticities of our farm sample are reported in Table 6. They have expected signs and, because of the crop disaggregation level of our data, are larger than those commonly found in the literature. The decomposition of these elasticities shows that a large part of the price effects on acreages can be due to the inclusion or not of these crops in the production regimes chosen by farmers. For crops like corn or pea, which are minor crops in the considered area, changes in the production regimes account for about one third of the estimated price elasticities. However, changes in the production regimes can also have significant effects for frequently produced crops. For instance, they account for 11% of the sugar beet acreage own price elasticities.

Table 6. Average Own Price Elasticities of Crop Acreages

	Winter wheat	*Corn*	*Spring barley*	*Sugar beet*	*Alfalfa*	*Protein pea*	*Rape- -seed*
Average crop acreage own price elasticities	0.33	4.26	0.44	1.39	0.74	1.22	0.76
Due to changes in acreages within production regimes	0.33 (100%)	2.33 (55%)	0.43 (98%)	1.24 (89%)	0.60 (81%)	0.71 (58%)	0.75 (99%)
Due to changes in production regimes	0.00 (0%)	1.93 (45%)	0.01 (2%)	0.15 (11%)	0.14 (19%)	0.51 (42%)	0.01 (1%)

Table 7. Per Regime Average Own Price Crop Acreage Elasticities

Regime			Crops produced in the regime						
Number	*Frequency*	*Crop number*	*Winter wheat*	*Corn*	*Spring barley*	*Sugar beet*	*Alfalfa*	*Peas*	*Rape--seed*
1	6.6%	7	0.33	0.95	0.92	1.19	0.62	0.68	0.84
2	21.5%	6	0.31		0.32	1.17	0.61	0.67	0.82
3	11.8%		0.32	0.95	0.92	1.16	0.57		0.75
4	48.6%	5	0.30		0.32	1.14	0.56		0.74
5	2.8%		0.31	0.95	0.90	1.10			0.44
6	2.5%	4	0.29	0.95	0.90				0.35
7	4.9%		0.29		0.31	1.10			0.43
8	1.3%	3	0.27		0.30				0.30

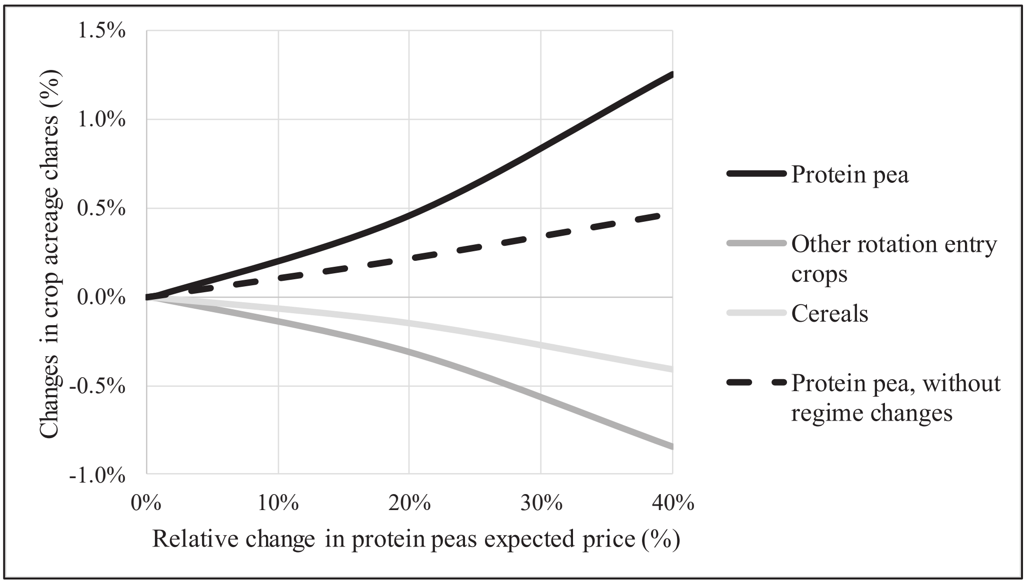
Observing how crop acreage elasticities within production regimes vary across regimes allows to illustrate the main features distinguishing ERS-MEMC models from their CR-MEMC counterparts. Table 7 reports the estimated means of own price crop acreage elasticities per regime.

These estimates display significant differences across production regimes. In particular, crop acreage own price elasticities grow with the number of crops produced in the considered production regime. The higher the crop number, the more farmers can make use of crop acreage substitution opportunities. For instance, the more the considered regime contains rotation starting crops, the more rapeseed acreage choices are responsive to rapeseed price. This elasticity range, on average, from 0.30, when rapeseed is the only rotation starting crop in the regime, to 0.84, in regimes with 4 rotation entry crops. Similarly, barley acreages are much more responsive to changes in barley price in regimes including corn than in regimes without corn. Corn and barley are the only spring cereals in farmers' crop set. Crop acreage models of CR-MEMC models cannot represent the substitution patterns uncovered by our estimation results. These models account for crop regimes but consider the same crop acreage model for all production regimes.

The impact of the production regime choice is further highlighted by simulating the effects of increases in the price of protein pea on its acreages. Owing to its fixing atmospheric nitrogen for themselves as well as for following crops, French agricultural scientists consider pea as a "diversification crop" of particular interest. Yet, protein pea acreages have declined over the last decade in the considered area mostly because of lacking profitability, especially as regards to that of other rotation starting crops. The simulated impacts of increases in the price of peas on crop acreages are depicted in Figure 3.

According to our results, a 40% increase in the expected price of pea would increase the average pea acreage share by 1.3%, from 2.0% to 3.3%. These additional pea acreages would mainly replace those of other rotation starting crops. The combined average acreage share of rapeseed, alfalfa and sugar beet would decrease by 0.9% while that of cereals would only decrease by 0.4%. This illustrates the interest in considering crop – agronomic and management – characteristics when specifying the acreage management cost function. Interestingly, about two thirds of the increase in the pea acreage would be due to new producers. This also explains another feature of our simulation results. The simulated increases in the pea acreage is not linear in the price of pea. In particular, the increase in the pea acreages is more pronounced above the 20% price increase level than below. Threshold effects due to production regime fixed costs and changes in crop acreage elasticities due to regime changes can explain this pattern. These induce kinks in farmers' pea acreage choices that are smoothed by the averaging process.

Figure 3. Estimated impacts of protein pea expected price on crop acreage shares



Source : Authors' calculations based on estimation results.

6. Concluding remarks

The main aims of this article are threefold. First, we present an original modelling framework for dealing with null acreages in MEMC models. This framework is fully consistent from an economic viewpoint and explicitly considers regime fixed costs. These features make the ERS-MEMC model proposed in this article suitable for analyzing and, to some extent, disentangling, the effects of the main drivers of farmers' acreage choices at disaggregation levels at which issues raised by null acreages are pervasive. Our estimation and simulation results notably tend to demonstrate that expected crop returns are not the sole significant drivers of farmers' crop acreage choices, at least in the short run. In particular, crop production fixed costs also matter. These results also show that crop acreages display patterns that cannot be accounted for by the CR-MCEM models currently used for handling null acreage choices. Effects of economic incentives on the crop acreage choices of a farmer strongly depend on the crop set chosen by the considered farmer.

Second, the application presented in this article illustrates the empirical tractability of random parameter ERS-MEMC models for investigating farmers' production choices. Of course, estimating such models raises challenging issues. But, this is also necessary for estimating structured micro-econometric models suitably accounting for important features characterizing micro-economic agricultural production data, among which significant unobserved heterogeneity. In particular, to estimate such models enables analysts to calibrate simulation models consisting of samples of farm specific models.

Third, according to our experience, ML estimators computed with stochastic versions of EM algorithms appear to be interesting alternatives to Simulated ML estimators for relatively large systems of interrelated equations such as the random parameter ERS-MEMC models considered in our empirical application. SAEM algorithms appear to be particularly relevant.

Of course, significant specification and estimation issues remain to be addressed. First, the empirical tractability of the ERS-MEMC model proposed in this article strongly relies on properties that are specific to the MNL crop acreage share models proposed by Carpentier and Letort (2014). Adapting our modelling approach to other crop acreage choice models would widen the scope of specification search for ERS-MEMC models. Also, the ERS-MEMC model considered in our empirical application relies on restrictive assumptions regarding the dynamic features of multi-crop technologies and farmers' choice process. Finally, the estimation cost of the models proposed in this article is relatively high, due to long computing and coding times.

This article proposes solutions to methodological issues that could be used for improving micro-econometric analyzes of policies impacting crop acreage choices. For instance, Babcock (2015) noted that policies related to biofuels led to a dramatic increase in interest in the econometric analyses of crop supply response to crop prices. The ERS-MEMC models considered in this article not only allow to disentangle intensive and extensive margin effects, they also allow to investigate crop choice effects. Analyzing crop choices also appear crucial for investigating agri-environmental policies and issues. For instance, changes in the location of crop production induced by climate change are due to crop set choices made the farm level. Also, as fostering crop diversification tend to become an important agri-environmental objective in many countries, including those of the European Union, coherent model of farmers' crop set choices appear to be especially relevant. Finally, random parameter ERS-MEMC models can contribute to close the gap existing between MEMC models and mathematical programming models (*e.g.*, Heckeleï *et al.*, 2012; Mérel and Howitt, 2014). The overall structure of our ERS-MEMC models is similar to that of mathematical programming models and their random parameter versions can be used for calibrating heterogeneous farm models.

Estimation costs appears to be among the limitations of our modelling framework that need to be addressed. Significant computing and coding costs make applied research work, such as specification search, tedious and time consuming. Relatively slight modifications of the model specification could, however, significantly reduce the estimation cost of the ERS-MEMC model presented in this article. For instance, in the model considered in the empirical application the covariance parameters of the random parameter vector $\boldsymbol{\gamma}_i$ represent more than 64% of the (786) estimated parameters. Yet, our estimates demonstrate that the random parameters of the crop yield supply and input demand models are strongly correlated, for a given crop but also across crops. This suggests that these parameters are linked by a few farmer specific "productivity factors". Such latent productivity factors could be used for imposing some structure on the variance-covariance matrix of the considered random parameters and, thereby, for significantly reducing the number of covariance parameters to be estimated. Also, relying on Laplace approximates for the regime choice probability function of the ESR-MEMC model involves tedious and time consuming computations. Less accurate but significantly simpler approximation approaches could dramatically reduce the estimation burden. But, these cruder approximation approaches could also impact the consistency of the estimated model. The extent of these impacts and the related – estimation burden *versus* specification approximation – trade-off is worth investigating in future work.

Références

Arndt, C. (1999). “Demand for herbicide in corn: an entropy approach using micro-level data.” *Journal of Agricultural and Resource Economics* 24(1): 204-221.

Arndt, C., Liu, S., and P.V. Preckel. (1999). “On dual approaches to demand systems estimation in the presence of binding quantity constraints.” *Applied Economics* 31(8): 999-1008.

Babcock, B.A. 2015. “Extensive and intensive agricultural supply response.” *Annual Review of Resources Economics* 7: 333-348.

Bayramoglu, B., and R. Chakir. (2016). “The impact of high crop prices on the use of agro-chemical inputs in France: A structural econometric analysis.” *Land Use Policy* 55: 204-211.

Bowman, M.S., and D. Zilbermann. (2013). “Economic factors affecting diversified farming.” *Ecology and Society* 18(1): 33.

Caffo, B.S., Jank, W., and G.L. Jones. (2005). “Ascent-based Monte Carlo expectation-maximization.” *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 67(2): 235-251.

Carpentier, A., and E. Letort. (2014). “Multicrop production models with Multinomial Logit acreage shares.” *Environmental and Resource Economics* 59(4): 537-559.

Carpentier, A., and E. Letort. (2012). “Accounting for heterogeneity in multicrop micro-econometric models: implications for variable input demand modeling.” *American Journal of Agricultural Economics* 94(1): 209-224.

Chakir, R., and A. Thomas. (2003). “Simulated maximum likelihood estimation of demand systems with corner solutions and panel data application to industrial energy demand.” *Revue d’Economie Politique* 113(6): 773-799.

Chambers, R., and R. Just. (1989). “Estimating Multiouput Technologies.” *American Journal of Agricultural Economics* 71(4): 980-95.

Chavas, J.-P., and M.T. Holt. (1990). “Acreage decisions under risk: the case of corn and soybeans.” *American Journal of Agricultural Economics* 72(3): 529-538.

Coyle, B.T. (1992). “Risk Aversion and Price Risk in Duality Models of Production: A Linear Mean-Variance Approach.” *American Journal of Agricultural Economics* 74(4): 849-859.

Coyle, B.T. (1999). “Risk Aversion and Yield Uncertainty in Duality Models of Production: A Linear Mean-Variance Approach.” *American Journal of Agricultural Economics* 81(3): 553-567.

Delyon, B., Lavielle, M., and E. Moulines. (1999). “Convergence of a stochastic approximation version of the EM algorithm.” *Annals of Statistics* 27(1): 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)* 1-38.

Fezzi, C., and I.J. Bateman. (2011). "Structural agricultural land use modeling for spatial agro-environmental policy analysis." *American Journal of Agricultural Economics* 93(4): 1168-1188.

Guyomard, H., Baudry, M., and A. Carpentier. (1996). "Estimating crop supply response in the presence of farm programmes: application to the CAP." *European Review of Agricultural Economics* 23(4): 401-420.

Harding, M.C., and J. Hausman. (2007). "Using a Laplace approximation to estimate the random coefficients logit model by nonlinear least squares." *International Economic Review* 48(4): 1311-1328.

Heckelei, T., Britz, W., and Y. Zhang. (2012). "Positive mathematical programming approaches-recent developments in literature and applied modelling." *Bio-Based and Applied Economics* 1(1): 109-124.

Heckeleï, T., and H. Wolff. (2003). "Estimation of constrained optimisation models for agricultural supply analysis based on generalised maximum entropy." *European Review of Agricultural Economics* 30(1): 27-50.

Holt, M.T. (1999). "A Linear Approximate Acreage Allocation Model." *Journal of Agricultural and Resource Economics* 24: 383-397.

Just, R.E., Zilberman, D., and E. Hochman. (1983). "Estimation of Multicrop Production Functions." *American Journal of Agricultural Economics* 65(4): 770-780.

Just, R.E., Zilberman, D., Hochman, E., and Z. Bar-Shira. (1990). "Input Allocation in Multicrop Systems." *American Journal of Agricultural Economics* 72(1): 200-209.

Kao, C., Lee, L., and M.M. Pitt. (2001). "Simulated maximum likelihood estimation of the linear expenditure system with binding non-negativity constraints." *Annals of Economics and Finance* 2(1): 203-223.

Keane, M.P. (2009). "Simulated maximum likelihood estimation based on first-order conditions." *International Economic Review* 50, 627-675.

Koutchadé, O.P., Carpentier, A., and F. Femenia. (2018). "Modeling heterogeneous farm responses to EU biofuel support with a random parameter multi-crop model." *American Journal of Agricultural Economics* 100(2): 434-455.

Kremen, C., A. Iles and C. Bacon. (2012). "Diversified farming systems: an agroecological, systems-based alternative to modern industrial agriculture." *Ecology and Society* 17(4): 44.

Kuhn, E., and M. Lavielle. (2005). “Maximum likelihood estimation in nonlinear mixed effects models.” *Computational Statistics and Data Analysis* 49(4): 1020-1038.

Lacroix, A., and A. Thomas. (2011). “Estimating the environmental impact of land and production decisions with multivariate selection rules and panel data.” *American Journal of Agricultural Economics* 93(3): 784-802.

Lavielle, M. (2014). *Mixed effects models for the population approach: models, tasks, methods and tools*. New York: Chapman and Hall/CRC.

Lee, L.F., and M.M. Pitt. (1987). “Microeconometric models of rationing, imperfect markets, and non-negativity constraints.” *Journal of Econometrics* 36(1-2): 89-110.

Lee, L.F., and M.M. Pitt. (1986). “Microeconometric demand system with binding nonnegativity constraints: the dual approach.” *Econometrica* 54(5): 1237-1242.

Lin, B.B. (2011). “Crop Diversification: Adaptative Management for Environmental Change.” *BioScience* 61(3): 183-193.

Matson, P.A., Parton, W.J., Power, A.G., and M.J. Swift. (1997). “Agricultural Intensification and Ecosystem Properties.” *Science* 277: 504-508.

McDonald, J.F., and R.A. Moffitt. (1980). “The uses of Tobit analysis.” *The Review of Economics and Statistics* 62(2): 318-321.

McLachlan, G., and T. Krishnan. (2007). *The EM algorithm and extensions*, 2nd. ed. New York: John Wiley & Sons.

Mérel, P., and R. Howitt. (2014). “Theory and application of positive mathematical programming in agriculture and the environment.” *Annual Review of Resources Economics* 6(1): 451-470.

Millimet, D.L., and R. Tchernis. (2008). “Estimating high-dimensional demand systems in the presence of many binding non-negativity constraints.” *Journal of Econometrics* 147(2): 384-385.

Moore, M., Gollehon, N. R., and M. B. Carey. (1994). “Multicrop Production Decisions in Western Irrigated Agriculture: The Role of Water Price.” *American Journal of Agricultural Economics* 76(4): 385-395.

Moore, M.R., and D.H. Negri. (1992). “A multicrop production model of irrigated agriculture, applied to water allocation policy of the Bureau of Reclamation.” *Journal of Agricultural and Resource Economics* 17(1): 29-43.

Oude Lansink, A. (2008). “Area allocation under price uncertainty on Dutch arable farms.” *Journal of Agricultural Economics* 50(1): 93-105.

Oude Lansink, A., and J. Peerlings. (1996). “Modelling the new EU cereals and oilseeds regime in the Netherlands.” *European Review of Agricultural Economics* 23(2): 161-178.

Ozarem, P.F., and J.A. Miranowski. (1994). “A Dynamic Model of Acreage Allocation with General and Crop-Specific Soil Capital.” *American Journal of Agricultural Economics* 76(3): 385-395.

Platoni, S., Sckokai, P., and D. Moro. (2012). “Panel data estimation techniques and farm-level data models.” *American Journal of Agricultural Economics* 94(5): 1202-1217.

Pudney, S. (1989). *Modelling Individual Choice: the Econometrics of Corners, Kinks and Holes.* Oxford: Basil Blackwell.

Ruud, P.A. (1991). “Extensions of estimation methods using the EM algorithm.” *Journal of Econometrics* 49(3): 305-341.

Serra, T., Zilberman, D., Goodwin, B.K., and K. Hyvonen. (2005). “Replacement of agricultural price supports by area payments in the European Union and the effects on pesticide use.” *American Journal of Agricultural Economics* 87(4): 870-884.

Shonkwiler, J.S., and S.T. Yen. (1999). “Two-Step Estimation of a Censored System of Equations.” *American Journal of Agricultural Economics* 81(3): 972-982.

Sckokai, P., and D. Moro. (2006). “Modeling the reforms of the common agricultural policy for arable crops under uncertainty.” *American Journal of Agricultural Economics* 88(1): 43-56.

Sckokai, P., and D. Moro. (2009). “Modelling the impact of the CAP Single Farm Payment on farm investment and output.” *European Review of Agricultural Economics* 36(2): 395-423.

Stanfield, P.M., Wilson, J.R., Mirka, G.A., Glasscock, N.F., Psihogios, J.P., and J.R. Davis. (1996). “Multivariate input modeling with Johnson distributions.” *Proceedings of the 28th conference on Winter simulation, IEEE Computer Society*, pp. 1457-1464.

Tilman, D., Cassman, K.G., Matson, P.A., Naylor, R. and S. Polasky. (2002). “Agricultural sustainability and intensive production practices.” *Nature* 418: 671-677.

Wales, T.J., and, A.D. Woodland. (1983). “Estimation of consumer demand systems with binding non-negativity constraints.” *Journal of Econometrics* 21(3): 263-285.

Wei, G.C., 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.

Wu, C.J. (1983). “On the convergence properties of the EM algorithm.” *The Annals of Statistics* 11(1), 95-103.

Appendix. SAEM algorithm structure

The aim of our estimation procedure is to compute the ML estimator of $\boldsymbol{\theta}_0$ or, at least, an estimator that is asymptotically equivalent to this estimator which is practically "infeasible". The ML estimator of $\boldsymbol{\theta}_0$ is the solution in $\boldsymbol{\Theta}$ to the ML problem $\max_{\boldsymbol{\theta}} L_N(\boldsymbol{\theta})$ where $L_N(\boldsymbol{\theta}) = \sum_{i=1}^{N} \ln \ell_i(\boldsymbol{\theta}) = \sum_{i=1}^{N} \int \left(\prod_{t=1}^{T} f(\mathbf{w}_{it}^{+}, r_{it} \,|\, \mathbf{z}_{it}, \boldsymbol{\gamma}; \boldsymbol{\delta}, \boldsymbol{\Psi}) \right) \varphi(\boldsymbol{\gamma} - \boldsymbol{\mu}; \boldsymbol{\Omega}) d\boldsymbol{\gamma}$. EM algorithms were proposed by (Dempster *et al*, 1977) for maximizing the likelihood function of models involving missing information, of which random parameter models are typical examples. Extensions of the original EM algorithm were then proposed for overcoming limitations of this algorithm (e.g., McLachlan and Krishnan, 2007; Lavielle, 2014), including issues such as those raised by the integration of the likelihood function of our model.

EM type algorithms are constructed based on the expectation conditional on the "observed data" of the "complete data" sample log-likelihood function of the considered model. Contribution of farmer *i* to the likelihood function of the model corresponds to the pdf of her/his sequence of production choices conditional on the sequence of exogenous variables characterizing this choice sequence. This choice sequence is given by $(\mathbf{w}_i^{+}, \mathbf{r}_i)$ with $\mathbf{w}_i^{+} = (\mathbf{w}_{it}^{+} : t = 1, ..., T)$ and $\mathbf{r}_i = (r_{it} : t = 1, ..., T)$ and the corresponding conditioning set by $\mathbf{z}_i \equiv (\mathbf{z}_{it} : t = 1, ..., T)$. The complete – observed and unobserved modelled variables – data related to farmer *i* consist of her/his observed production choice sequence, $(\mathbf{w}_i^{+}, \mathbf{r}_i)$, and her/his specific parameter vector, $\boldsymbol{\gamma}_i$. The complete data log-likelihood function of our model is thus given by:

$$\sum_{i=1}^{N} \ln \ell_i^{C}(\boldsymbol{\theta}, \boldsymbol{\gamma}_i) = \sum_{i=1}^{N} \sum_{t=1}^{T} \ln f(\mathbf{w}_{it}^{+}, r_{it} \,|\, \mathbf{z}_{it}, \boldsymbol{\gamma}_i; \boldsymbol{\delta}, \boldsymbol{\Psi}) + \sum_{i=1}^{N} \ln \varphi(\boldsymbol{\gamma}_i; \boldsymbol{\mu}, \boldsymbol{\Omega}). \quad \text{(A.1)}$$

The "observed data" related to farmer *i*, thereafter denoted by $\boldsymbol{\kappa}_i$, consist of her/his observed production choice sequence, $(\mathbf{w}_i^{+}, \mathbf{r}_i)$, and of the exogenous variables conditioning these choice sequence, $\mathbf{z}_i$. That is to say, $\boldsymbol{\kappa}_i = (\mathbf{w}_i^{+}, \mathbf{r}_i, \mathbf{z}_i)$. According to our notations function $f(\boldsymbol{\gamma}_i \,|\, \boldsymbol{\kappa}_i; \boldsymbol{\theta}_0)$ denotes the density of $\boldsymbol{\gamma}_i$ conditional on $\boldsymbol{\kappa}_i$. Let function

$$E_{(\bar{\boldsymbol{\theta}})}[\ln \ell_i^{C}(\boldsymbol{\theta}, \boldsymbol{\gamma}_i) \,|\, \boldsymbol{\kappa}_i] = \int \ln \ell_i^{C}(\boldsymbol{\theta}, \boldsymbol{\gamma}_i) f(\boldsymbol{\gamma} \,|\, \boldsymbol{\kappa}_i; \bar{\boldsymbol{\theta}}) d\boldsymbol{\gamma} \quad \text{(A.2)}$$

denote the expectation of the $\ln \ell_i^{C}(\boldsymbol{\theta}, \boldsymbol{\gamma}_i)$ conditional on $\boldsymbol{\kappa}_i$ based on the pdf $f(\boldsymbol{\gamma}_i \,|\, \boldsymbol{\kappa}_i; \bar{\boldsymbol{\theta}})$ where $\bar{\boldsymbol{\theta}}$ is a candidate estimate of $\boldsymbol{\theta}_0$. Function

$$Q(\boldsymbol{\theta} \,|\, \bar{\boldsymbol{\theta}}) = \sum_{i=1}^{N} E_{(\bar{\boldsymbol{\theta}})}[\ln \ell_i^{C}(\boldsymbol{\theta}, \boldsymbol{\gamma}_i) \,|\, \boldsymbol{\kappa}_i] \quad \text{(A.3)}$$

defines the expectation of the complete data sample log-likelihood function $\sum_{i=1}^{N} \ln \ell_i^{C}(\boldsymbol{\theta}, \boldsymbol{\gamma}_i)$ conditional on $(\boldsymbol{\kappa}_i : i = 1, ..., N)$ based on the pdfs $f(\boldsymbol{\gamma}_i \,|\, \boldsymbol{\kappa}_i; \bar{\boldsymbol{\theta}})$ for $i = 1, ..., N$. This function, which can be interpreted as a well-behaved proxy of $L_N(\boldsymbol{\theta})$ when $\bar{\boldsymbol{\theta}}$ is suitably chosen, is the "engine" of EM type algorithms that can be used for estimating $\boldsymbol{\theta}_0$.

SAEM algorithms iterate three steps until numerical convergence: a Simulation (S) step, an Approximation (A) step and a Maximization (M) step. They generate sequences of estimates of $\boldsymbol{\theta}_0$ that converge to maxima of $L_N(\boldsymbol{\theta})$ under mild assumptions, thereby allowing to compute ML estimators of $\boldsymbol{\theta}_0$ (Delyon *et al*, 1999; Kuhn et Lavielle, 2005; Lavielle, 2014). Assuming that the estimate of $\boldsymbol{\theta}_0$ obtained at the end of iteration n is given by $\boldsymbol{\theta}_{(n)}$, our SAEM algorithm proceeds as follows at iteration $n+1$.

The S step consists of integrating terms $E_{(n)}[\ln \ell_i^C(\boldsymbol{\theta},\boldsymbol{\gamma}_i) \mid \boldsymbol{\kappa}_i] = \int \ln \ell_i^C(\boldsymbol{\theta},\boldsymbol{\gamma}_i) f(\boldsymbol{\gamma} \mid \boldsymbol{\kappa}_i;\boldsymbol{\theta}_{(n)}) d\boldsymbol{\gamma}$ with simulation methods for $i=1,...N$. Building on the work of Caffo *et al* (2005), we use an Importance Sampling approach. Terms $E_{(n)}[\ln \ell_i^C(\boldsymbol{\theta},\boldsymbol{\gamma}_i) \mid \boldsymbol{\kappa}_i]$ are approximated with simulated weighted sums $\sum_{j=1}^{J_{(n)}} \tilde{\varpi}_{i,(n)}^j \ln \ell_i^C(\boldsymbol{\theta},\tilde{\boldsymbol{\gamma}}_{i,(n)}^j)$ where terms $\tilde{\boldsymbol{\gamma}}_{i,(n)}^j$ are independent random draws from $\mathcal{N}(\boldsymbol{\mu}_{(n)},\boldsymbol{\Omega}_{(n)})$, while terms

$$\tilde{\varpi}_{i,(n)}^j = \frac{\prod_{t=1}^T f(\mathbf{w}_{it}^+, r_{it} \mid \mathbf{z}_{it}, \tilde{\boldsymbol{\gamma}}_{i,(n)}^j; \hat{\boldsymbol{\delta}}_{(n)}, \hat{\boldsymbol{\Psi}}_{(n)})}{\sum_{j=1}^{J_{(n)}} \prod_{t=1}^T f(\mathbf{w}_{it}^+, r_{it} \mid \mathbf{z}_{it}, \tilde{\boldsymbol{\gamma}}_{i,(n)}^j; \hat{\boldsymbol{\delta}}_{(n)}, \hat{\boldsymbol{\Psi}}_{(n)})} \tag{A.4}$$

are their corresponding normalized importance weights, for $j=1,...,J_{(n)}$. Other proposed densities are more efficient than that of $\mathcal{N}(\boldsymbol{\mu}_{(n)},\boldsymbol{\Omega}_{(n)})$ but are more difficult to draw from.

The A step consists of constructing function $\tilde{Q}_{(n)}(\boldsymbol{\theta})$, the stochastic approximation of $Q(\boldsymbol{\theta} \mid \boldsymbol{\theta}_{(n)})$, by using the following recursive formula:

$$\tilde{Q}_{(n)}(\boldsymbol{\theta}) = (1-\lambda_{(n)})\tilde{Q}_{(n-1)}(\boldsymbol{\theta}) + \lambda_{(n)} \sum_{i=1}^N \sum_{j=1}^{J_{(n)}} \tilde{\varpi}_{i,(n)}^j \ln \ell_i^C(\boldsymbol{\theta},\tilde{\boldsymbol{\gamma}}_{i,(n)}^j). \tag{A.5}$$

Delyon *et al* (1999) and Kuhn and Lavielle (2005) provide guidelines for suitably choosing the sequence of weight terms $\lambda_{(n)}$, which must lie in $(0,1]$. Large values of $\lambda_{(n)}$ allow to explore the parameter space and yield a quick convergence to the neighborhood of a solution to the ML problem. But they also imply large simulation noise. Reducing the value of $\lambda_{(n)}$ reduces the simulation noise and allow the algorithm to converge in the neighborhood of a solution to the ML problem.

Kuhn and Lavielle (2005) also provide guidelines for suitably choosing the number of draws $J_{(n)}$. Importantly, large numbers of random draws are not needed at each iteration since function $\tilde{Q}_{(n)}(\boldsymbol{\theta})$, by construction, reuses the random draws obtained in previous iterations. Indeed, "recycling" previous iteration draws is a major advantage of SAEM algorithms over their competing alternatives such as MCEM algorithms (Delyon *et al*, 1999; Lavielle, 2014).

As a matter of fact, the SAEM algorithm presented here performs significantly better that its MCEM counterpart, which corresponds to the SAEM algorithm with $\lambda_{(n)} = 1$, in our empirical application.

The M step consists of updating the estimate of $\boldsymbol{\theta}_0$ by computing $\boldsymbol{\theta}_{(n+1)}$. This updated estimate is defined either as:

$$\boldsymbol{\theta}_{(n+1)} = \arg\max_{\boldsymbol{\theta}} \tilde{Q}_{(n)}(\boldsymbol{\theta}) \tag{A.6}$$

or as any $\boldsymbol{\theta}_{(n+1)}$ such that condition $\tilde{Q}_{(n)}(\boldsymbol{\theta}_{(n+1)}) \geq \tilde{Q}_{(n)}(\boldsymbol{\theta}_{(n)})$ holds.

The main advantage of SAEM algorithms, and of other EM type algorithms, for maximizing the log-likelihood function of random parameters models is due to the following decomposition of $\tilde{Q}_{(n)}(\boldsymbol{\theta})$:

$$\tilde{Q}_{(n)}(\boldsymbol{\theta}) = \tilde{W}_{(n)}(\boldsymbol{\delta}, \boldsymbol{\Psi}) + \tilde{M}_{(n)}(\boldsymbol{\mu}, \boldsymbol{\Omega}) \tag{A.7}$$

where:

$$\tilde{W}_{(n)}(\boldsymbol{\delta}, \boldsymbol{\Psi}) = (1-\lambda_{(n)})\tilde{W}_{(n-1)}(\boldsymbol{\delta}, \boldsymbol{\Psi}) + \lambda_{(n)} \sum_{i=1}^{N} \sum_{j=1}^{J_{(n)}} \tilde{\omega}_{i,(n)}^{j} \sum_{t=1}^{T} \ln f(\mathbf{w}_{it}^{+}, r_{it} \mid \mathbf{z}_{it}, \tilde{\boldsymbol{\gamma}}_{i,(n)}^{j}; \boldsymbol{\delta}, \boldsymbol{\Psi}) \tag{A.8}$$

and:

$$\tilde{M}_{(n)}(\boldsymbol{\mu}, \boldsymbol{\Omega}) = (1-\lambda_{(n)})\tilde{M}_{(n-1)}(\boldsymbol{\mu}, \boldsymbol{\Omega}) + \lambda_{(n)} \sum_{i=1}^{N} \sum_{j}^{J_{(n)}} \tilde{\omega}_{i,(n)}^{j} \ln \varphi(\tilde{\boldsymbol{\gamma}}_{i,(n)}^{j} - \boldsymbol{\mu}; \boldsymbol{\Omega}) . \tag{A.9}$$

This decomposition enables us to separately update parameters $(\boldsymbol{\mu}, \boldsymbol{\Omega})$ and $(\boldsymbol{\delta}, \boldsymbol{\Psi})$. Moreover, term $(\boldsymbol{\mu}_{(n+1)}, \boldsymbol{\Omega}_{(n+1)}) = \arg\max_{\boldsymbol{\theta}} \tilde{M}_{(n)}(\boldsymbol{\mu}, \boldsymbol{\Omega})$ can be obtained in analytical closed form based on the "sufficient statistic approach" proposed by Delyon *et al* (1999).

Maximizing $\tilde{W}_{(n)}(\boldsymbol{\delta}, \boldsymbol{\Psi})$ in $(\boldsymbol{\delta}, \boldsymbol{\Psi})$ appears to be much more difficult due to the functional form of $f(\mathbf{w}_{it}^{+}, r_{it} \mid \mathbf{z}_{it}, \boldsymbol{\gamma}_i; \boldsymbol{\eta})$. Yet, function $\tilde{W}_{(n)}(\boldsymbol{\delta}, \boldsymbol{\Psi})$ can be rewritten as:

$$\tilde{W}_{(n)}(\boldsymbol{\delta}, \boldsymbol{\Psi}) = \tilde{W}_{(n)}^{yxs}(\boldsymbol{\delta}, \boldsymbol{\Psi}) + \tilde{W}_{(n)}^{r}(\boldsymbol{\delta}, \boldsymbol{\Psi}) \tag{A.10}$$

where:

$$\tilde{W}_{(n)}^{yxs}(\boldsymbol{\delta}, \boldsymbol{\Psi}) = (1-\lambda_{(n)})\tilde{W}_{(n-1)}^{yxs}(\boldsymbol{\delta}, \boldsymbol{\Psi}) + \lambda_{(n)} \sum_{i=1}^{N} \sum_{j=1}^{J_{(n+1)}} \tilde{\omega}_{i,(n)}^{j} \sum_{t=1}^{T} \ln f(\mathbf{w}_{it}^{+} \mid \mathbf{z}_{it}, \tilde{\boldsymbol{\gamma}}_{i,(n)}^{j}; \boldsymbol{\delta}, \boldsymbol{\Psi}) \tag{A.11}$$

and:

$$\tilde{W}_{(n)}^{r}(\boldsymbol{\delta}, \boldsymbol{\Psi}) = (1-\lambda_{(n)})\tilde{W}_{(n-1)}^{r}(\boldsymbol{\delta}, \boldsymbol{\Psi}) + \lambda_{(n)} \sum_{i=1}^{N} \sum_{j=1}^{J_{(n+1)}} \tilde{\omega}_{i,(n)}^{j} \sum_{t=1}^{T} P(r_{it} \mid \mathbf{z}_{it}, \tilde{\boldsymbol{\gamma}}_{i,(n)}^{j}, \mathbf{s}_{it}^{+}; \boldsymbol{\delta}, \boldsymbol{\Psi}) . \tag{A.12}$$

Terms $\ln f(\mathbf{w}_{it}^{+} \mid \mathbf{z}_{it}, \boldsymbol{\gamma}_i; \boldsymbol{\delta}, \boldsymbol{\Psi})$ are defined – up to an additive term that doesn't depend on $(\boldsymbol{\delta}, \boldsymbol{\Psi})$ – as log-likelihood functions at $(\boldsymbol{\delta}, \boldsymbol{\Psi})$ of a Gaussian Seemingly Unrelated (linear) Regression (SUR) system with dependent variables missing at random (conditionally on $(\mathbf{z}_{it}, \boldsymbol{\gamma}_i)$).

Ruud (1991) discussed the use of EM algorithms for computing ML estimators of models with latent Gaussian SUR systems. Based on Ruud's insights, we devised a simple EM type procedure aimed at obtaining values of $(\boldsymbol{\delta},\boldsymbol{\Psi})$ ensuring that condition $\tilde{W}_{(n)}^{yxs}(\boldsymbol{\delta},\boldsymbol{\Psi}) \geq \tilde{W}_{(n)}^{yxs}(\boldsymbol{\delta}_{(n)},\boldsymbol{\Psi}_{(n)})$ holds. This procedure defines $(\boldsymbol{\delta}_{(n+1)},\boldsymbol{\Psi}_{(n+1)})$, which, together with $(\boldsymbol{\mu}_{(n+1)},\boldsymbol{\Omega}_{(n+1)})$, completes $\boldsymbol{\theta}_{(n+1)}$.

We don't consider function $\tilde{W}_{(n)}^{r}(\boldsymbol{\delta},\boldsymbol{\Psi})$ when updating the estimate of $(\boldsymbol{\delta}_0,\boldsymbol{\Psi}_0)$ because this function is an awkward function of $(\boldsymbol{\delta},\boldsymbol{\Psi})$. It involves the regime choice probability functions $P(r_{it} \mid \mathbf{z}_{it},\boldsymbol{\gamma}_i,\mathbf{s}_{it}^{+};\boldsymbol{\delta},\boldsymbol{\Psi})$. Yet, our ignoring $\tilde{W}_{(n)}^{r}(\boldsymbol{\delta},\boldsymbol{\Psi})$ in the construction of $(\boldsymbol{\delta}_{(n+1)},\boldsymbol{\Psi}_{(n+1)})$ doesn't ensure that condition $\tilde{Q}_{(n)}(\boldsymbol{\theta}_{(n+1)}) \geq \tilde{Q}_{(n)}(\boldsymbol{\theta}_{(n)})$ holds whereas it is necessary for the convergence of the SAEM algorithm. We devised a simple heuristic for coping with cases where $\boldsymbol{\theta}_{(n+1)}$ doesn't succeed in increasing $\tilde{Q}_{(n)}(\boldsymbol{\theta})$ from $\tilde{Q}_{(n)}(\boldsymbol{\theta}_{(n)})$. Yet, this heuristic was rarely activated when running this SAEM algorithm for estimating the ESR-MEMC model considered in our application. Two explanations can be put forward. If regime choice probability functions $P(r_{it} \mid \mathbf{z}_{it},\boldsymbol{\gamma}_i,\mathbf{s}_{it}^{+};\boldsymbol{\delta},\boldsymbol{\Psi})$ don't have any "active" role when computing $(\boldsymbol{\delta}_{(n+1)},\boldsymbol{\Psi}_{(n+1)})$, they have important "passive" roles though their effects as elements of IS weights $\tilde{\omega}_{i,(n)}^{j}$. Also, the recursive structure of the considered ERS-MEMC model implies that most statistical information needed to estimate $(\boldsymbol{\delta}_0,\boldsymbol{\Psi}_0)$ is contained in farmers' crop level choices that are considered in $\tilde{W}_{(n)}^{yxs}(\boldsymbol{\delta},\boldsymbol{\Psi})$. Parameter $(\boldsymbol{\delta}_0,\boldsymbol{\Psi}_0)$ only impacts regime choices through its effects on the expected crop profitability levels.

Les **Working Papers SMART – LERECO** sont produits par l'UMR SMART-LERECO

- **UMR SMART-LERECO**

L'Unité Mixte de Recherche (UMR 1302) *Laboratoire d'Etudes et de Recherches en Economie sur les Structures et Marchés Agricoles, Ressources et Territoires* comprend les unités de recherche en Economie INRAE de Rennes, INRAE de Nantes et les membres des Unités Pédagogiques de Rennes et Angers du département Economie, Gestion et Société d'Agrocampus Ouest.

Adresse:
UMR SMART-LERECO, 4 allée Adolphe Bobierre, CS 61103, 35011 Rennes cedex

Site internet : https://www6.rennes.inrae.fr/smart

Liste complète des Working Papers SMART – LERECO :
https://www6.rennes.inrae.fr/smart/Working-Papers
https://ideas.repec.org/s/rae/wpaper.html
http://ageconsearch.umn.edu/handle/204962/

The **Working Papers SMART – LERECO** are produced by UMR SMART-LERECO

- **UMR SMART-LERECO**

The « Mixed Unit of Research » (UMR1302) *Laboratory for Empirical Research in Economics on Structures and Markets in Agriculture, Resources and Territories* is composed of the research units in Economics of INRAE Rennes and INRAE Nantes and of the members of the Agrocampus Ouest's Department of Economics, Management and Society who are located in Rennes and Angers.

Address:
UMR SMART-LERECO, 4 allée Adolphe Bobierre, CS 61103, 35011 Rennes cedex

Website: https://www6.rennes.inrae.fr/smart_eng/

Full list of the Working Papers SMART – LERECO:
https://www6.rennes.inrae.fr/smart/Working-Papers
https://ideas.repec.org/s/rae/wpaper.html
http://ageconsearch.umn.edu/handle/204962/

Contact

Working Papers SMART – LERECO
INRAE, UMR SMART-LERECO
4 allée Adolphe Bobierre, CS 61103
35011 Rennes cedex, France
Email : smart-lereco-wp@inrae.fr

2020

Working Papers SMART – LERECO

UMR INRAE-Agrocampus Ouest **SMART-LERECO** (Laboratoire d'Etudes et de Recherches en Economie sur les Structures et Marchés Agricoles, Ressources et Territoires)

Rennes, France