Accounting for Unobserved Heterogeneity in Micro-Econometric Agricultural Production Models: A random Parameter Approach

Philippe Koutchade[1], Alain Carpentier[1, 2], and Fabienne Femenia[1]

[1]INRA UMR 1302 SMART
[2]Agrocampus Ouest UMR 1302 SMART

November, 2014

Abstract.

In this paper we rely on a random parameter approach to account for the unobserved heterogeneity of farms/farmers in the estimation of production choices model. We use a Stochastic Expectation-Maximization algorithm to estimate a Nested Multinomial Logit model of yield and acreage choices and perform some simulations using this model. This approach allows accounting for farms' and farmers' heterogeneity in a flexible way. Our results show that heterogeneity significantly matters in agricultural production choice models and that ignoring the heterogeneous determinants of farmers' choices can have important impacts on estimation and simulation outcomes of micro econometric models. Estimates of random parameter models such as the one presented here can be used for, at least, two purposes: for the calibration of simulation models accounting for farm unobserved heterogeneity and to investigate the potential explanations of this unobserved heterogeneity.

Keywords: heterogeneity, random parameter models, agricultural production choices
JEL codes: Q12, C13, C15

Introduction

Evidences of the effects of unobserved heterogeneity in micro-econometric models are now pervasive in many applied economics fields. During the last two decades applied micro-econometricians have developed tools to estimate models explicitly accounting for the effects of unobserved heterogeneity on economic choices. These tools have already been successfully used in several applied economics domains. Empirical studies highlighting the role of unobserved heterogeneity effects in econometric models can be found in labor economics, in industrial organization, in transportation economics or in international trade economics.

Consumer choice models assume that consumers' preferences are heterogeneous and that this unobserved heterogeneity has important effects on product demands, especially for differentiated products (see, *e.g.*, Ackerberg *et al* 2007). Wage equations are specified accounting for the fact that workers' abilities are heterogeneous. While unobserved, this heterogeneity in abilities is shown to determine a large part of the observed wage variability as well as of the returns to education (see, *e.g.*, Heckman and Sedlacek, 1985, Heckman, 2001). Firm choice models account for the fact that firms use different technologies and, as a consequence, that they have different productivity levels and various supply choices (see, *e.g.*, Eaton *et al* 2011). An important point is that the effects of unobserved heterogeneity are not simply added to the models considered above. These effects also affect the responses of these models to important interest variables.

Our view is that similar heterogeneity features characterize agricultural production choices. Farms and farmers are heterogeneous and this heterogeneity affects the way farmers respond to, *e.g.*, economic incentives.

The objectives of this article are twofold. First, we aim at showing that unobserved heterogeneity effects significantly matter in empirical agricultural production choice models. Second, we aim at showing that tools recently developed by micro-econometricians and statisticians allow specification and estimation of econometric agricultural production choice models accounting for farms' and farmers' unobserved heterogeneity in a fairly flexible way.

Farmers face different production conditions due to heterogeneous soil quality or usual climatic conditions across space. They also own different machineries and different wealth levels. Finally, farmers are also different because of their various educational level or abilities, as well as because they may have different objectives with respect to income risk or with respect to the leisure *versus* labor trade-off. These heterogeneity sources are likely to have important impacts on farmers' production choices.

But to control for the effects of these heterogeneity sources is difficult in practice, for two main reasons. First, as shown by the short list given above, potential heterogeneity sources are numerous. Second, many heterogeneity sources are not suitably described in the data sets usually used by agricultural economists. As a result, empirical investigators generally rely on a few variables – *e.g*. farms' size, farmers' age, farmers' education, farms' location or, when available, rough soil quality indices – to control for the effects of many heterogeneity sources on farmers' production choices. As a matter of fact, numerous important heterogeneity sources are unobserved for agricultural production modeling.

Means usually employed by agricultural production economists to cope with the unobserved heterogeneity of farms and farmers depend on their modeling approaches and purposes.

Mathematical programming models used to analyze agricultural supply responses to economic policies (or other determinants of farmers' choices) are usually built by considering sets of farms, of small regions or of farm-types. A mathematical programming model is calibrated for each element of the considered set of "farms". This disaggregated calibration procedure allows controlling for farms' and farmers' unobserved heterogeneity. Of course the lack of statistical background of the standard calibration procedures is often pointed out as a major limitation of agricultural supply mathematical programming models (Howitt 1995 ; Heckeleï and Wolff 2003 ; Heckeleï *et al* 2012). However, the simulations provided by these models appear to be highly valued by decision-makers. These provide disaggregated results with respect to the simulated effects of agricultural policy measures on farmers' choices across more or less large geographical areas.

By comparison, the ability of micro-econometric models of agricultural production choices to account for farms' and farmers' heterogeneity is much more limited. As recalled above, only a few control variables are usually available to agricultural production economists. Standard specifications of econometric agricultural production choice models can be defined as a sum of a deterministic part and of a vector of random error terms. In these models, farmers' responses to economic (or other) incentives are governed by the deterministic part – *i.e.* by a few statistically estimated parameters – and the effects of farms' and farmers' unobserved heterogeneity are "pushed" into additively separable error terms. This often leads to simulation results which are unrealistically homogeneous across farms.

The agricultural production choice models proposed in this article allow accounting for farms' and farmers' unobserved heterogeneity while being empirically tractable. They can also be used to design simulation models in which a parameter vector is "statistically calibrated" for each sampled farmer.

We adopt the random parameter modeling framework. This framework allows estimating standard production choice models under the assumption that the model parameters are farmer specific. Basically, the considered models allow the model of each sampled farmer to have its own parameter vector and, thereby, permit to account for unobserved heterogeneity effects across farmers. Standard data set, even panel data sets, do not permit direct estimation of the individual parameters. The objective of the estimation is to characterize the distribution of the model parameters across the considered farmer population.

We illustrate these points through the specification and the estimation of a multicrop econometric model with random parameters. The Multinomial Logit (MNL) framework proposed by Carpentier and Letort (2013) was chosen due to its simplicity, to its parameter parsimony and to the easy interpretation of its parameters. The specified model being parametric, we rely on the Maximum Likelihood (ML) framework for its estimation. More specifically we use estimators and optimization procedures specifically designed by statisticians for the estimation of a class of models to which random parameter models belong.

The empirical application considers a sample of French crop producers observed from 2004 to 2007. Obtained results demonstrate that unobserved heterogeneity matters for the modeling of micro-economic agricultural production choices, even within a small area. Key parameters of farmers' choice models are significantly affected by unobserved heterogeneity effects, *i.e.* exhibit significant variability across farmers. We also show how random parameter models can be used to "statistically calibrate" a multicrop simulation model based on a sample of heterogeneous farms. Simulation results show that it is important for the estimated models to allow farmers to respond heterogeneously to homogeneous economic incentives.

The general features of random parameter models are presented in the first section, together with their main advantages and limitations. The second section presents the multicrop econometric model that we consider in order to investigate the advantages of accounting for unobserved heterogeneity in agricultural production choice models. Identification and estimation issues are discussed in the third section. The estimation results and their interpretations are provided in the fourth section.

1. Unobserved heterogeneity and random parameter models

This section presents the main features of random parameter models of agricultural production choices. It also introduces important elements to be used in the presentation of the estimation issues. We consider short run production choices of farmers –*i.e.* an acreage (share) demand system and a yield supply system in the empirical application – and we take for granted that farmers' choices rely on heterogeneous determinants. We consider the use of panel data so that observations are indexed by $i = 1,...,N$ (farm/farmers) and $t = 1,...,T$ (year).

A random parameter model is composed of two parts. The first part of the model, the "behavioral model" (or the "kernel model" in Train's (2007, 2008) terminology), formally describes the causal process of interest and defines its statistical characteristics conditional on the considered random parameters (and on the exogenous variables). Basically, the "behavioral model" for agricultural production choices is a standard agricultural production choice model in which some or all parameters are chosen to be farmer specific, which imposes an examination of the statistical relationships between these parameters and the other elements of the model, *i.e.* its explanatory variables and error terms.

The second part of the model defines the characteristics of the distribution of the random parameters (conditional on the exogenous variables), *i.e.* the "mixing" distribution of the model according to the terminology used in statistics.

The equation

(1) $\mathbf{c}_{it} = \mathbf{r}(\mathbf{z}_{it}, \mathbf{e}_{it}; \mathbf{q}_i)$

describes the production choices $\mathbf{c}_{it}$ of farmer *i* in year *t* as a known response function **r** to the determinants of these choices $(\mathbf{z}_{it}, \mathbf{e}_{it})$, whether these determinants are observed or not. The term $\mathbf{z}_{it}$, respectively $\mathbf{e}_{it}$, is observed, respectively unobserved. The response function **r** is parameterized by a farmer specific parameter vector $\mathbf{q}_i$. Note that if the random terms $(\mathbf{e}_{it}, \mathbf{q}_i)$ are unobserved to the econometrician, they are known to farmer *i* in *t* and partly determine its choices through their effects in **r**. Equation (1) is a deep structural model, or an "all causes" model. It defines how the choice of farmer *i*, $\mathbf{c}_{it}$, is caused by its determinants $(\mathbf{z}_{it}, \mathbf{e}_{it})$ up the characteristics of this farmer and of his farm, $\mathbf{q}_i$. Equation (1) can be any agricultural production choice model where the usual fixed parameters, at least some of them, are replaced by farmer specific parameter vector $\mathbf{q}_i$.[1]

In a short run production choice context, the random parameters $\mathbf{q}_i$ mainly capture the effects of the farms' natural or quasi-fixed factor endowments, of the farmers' production technologies and of farmers' characteristics. The more these effects vary across farms in the considered population, the more likely is the distribution of $\mathbf{q}_i$ to exhibit significant variability.

Equation (1) is completed by statistical assumptions in order to define the "behavioral model" of the considered random parameter model. It is assumed here that $\mathbf{z}_{it}$ and $\mathbf{e}_{it}$ are independent conditionally on $\mathbf{q}_i$. *I.e.* it is assumed that controlling for the farms' and farmers' characteristics ensures that $\mathbf{z}_{it}$ can be interpreted as purely exogenous factors – such as market prices or climatic events – affecting farmers' choices. It is further assumed that $\mathbf{q}_i$ and $\mathbf{e}_{it}$ are statistically independent. This assumption relies on the idea that $\mathbf{q}_i$ captures the permanent unobserved characteristics of farmer *i* affecting $\mathbf{z}_{it}$ while $\mathbf{e}_{it}$ mostly represents the effects of idiosyncratic shocks on $\mathbf{z}_{it}$ – *i.e.* $\mathbf{e}_{it}$ basically is a "standard" error term.

[1] The functional form of **r** can also be defined up to a fixed parameter vector to be estimated, as in our empirical application. The response function can also depend on available variables describing the farmers or/and or their farms (these may also be included in the vector $\mathbf{z}_{it}$). These extensions are straightforward and are ignored here to reduce the notational burden.

We assume here that $\mathbf{c}_{it}$ (and thus $\mathbf{e}_{it}$) and $\mathbf{q}_i$ are continuous random variables. The behavioral model is parametric, as it is the case in our empirical application, if the distribution of $\mathbf{e}_{it}$, denoted by $\mathcal{D}(\mathbf{e}_{it})$, is assumed to be a member of a given parametric distribution family. This family is characterized by the density $g(\mathbf{e}_t;\boldsymbol{\mu}_t)$. Of course $\mathcal{D}(\mathbf{e}_{it})$ and equation (1) directly define the distribution of $\mathbf{c}_{it}$ conditional on $(\mathbf{z}_{it},\mathbf{q}_i)$, *i.e.* $\mathcal{D}(\mathbf{c}_{it}\,|\,\mathbf{z}_{it},\mathbf{q}_i)$. Assuming that the response function $\mathbf{r}$ is invertible in $\mathbf{e}_{it}$, the density of $\mathcal{D}(\mathbf{c}_{it}\,|\,\mathbf{z}_{it},\mathbf{q}_i)$ is given by:

(2) $$f(\mathbf{c}_{it}\,|\,\mathbf{z}_{it},\mathbf{q}_i;\boldsymbol{\mu}_t)=\left|\det\left(\tfrac{\partial}{\partial\mathbf{e}_{it}'}\mathbf{r}\left(\mathbf{z}_{it},\mathbf{r}^{(-1)}(\mathbf{z}_{it},\mathbf{c}_{it};\mathbf{q}_i);\mathbf{q}_i\right)\right)\right|^{-1} g\left(\mathbf{r}^{(-1)}(\mathbf{z}_{it},\mathbf{c}_{it};\mathbf{q}_i);\boldsymbol{\mu}_t\right).$$

Equation (1) and the independence assumptions described above define a "behavioral model" which can be used with cross-section data. With panel data additional assumptions are required in order to describe the eventual dynamic features of the considered choices. It is assumed here that the modeled choice process is essentially static in the sense that $\mathbf{z}_{it}$ and $\mathbf{e}_{it}$ are independent conditionally on $\mathbf{q}_i$ for any pair of years (s,t), *i.e.* $\mathbf{z}_{it}$ is assumed to be weakly exogenous with respect to $\mathbf{e}_{it}$ according to the panel data econometrics terminology. This condition simplifies the exposition and is assumed to hold in our empirical application.

This application considers short-run crop production choices and relies on a short panel data set, *i.e.* with $T=4$. These choices are modeled as static choices because the main dynamic aspects of crop production choices are due to crop rotations. Such dynamic effects can be suitably approximated by farmer specific parameters because crop rotation effects imply highly persistent dynamic effects in the crop production choices when farmers' base their production choices on a few rotation schemes. These elements also provide arguments for assuming that the $\mathbf{e}_{it}$ terms are independent across *t*.

Under these assumptions the joint density of the vector $\mathbf{c}_i\equiv(\mathbf{c}_{i1},...,\mathbf{c}_{iT})$ conditional on $(\mathbf{q}_i,\mathbf{z}_i)\equiv(\mathbf{q}_i,\mathbf{z}_{i1},...,\mathbf{z}_{iT})$ is given by:

(3) $$f(\mathbf{c}_i\,|\,\mathbf{z}_i,\mathbf{q}_i;\boldsymbol{\mu})=\prod\nolimits_{t=1}^{T} f(\mathbf{c}_{it}\,|\,\mathbf{z}_{it},\mathbf{q}_i;\boldsymbol{\mu}_t).$$

where $\boldsymbol{\mu} \equiv (\boldsymbol{\mu}_1 ..., \boldsymbol{\mu}_T)$. Of course, farmers' choices are linked across time due to their relying on the same parameter vector $\mathbf{q}_i$. But these choices are assumed to be independent across time conditionally on $\mathbf{q}_i$.

According to the assumption set describe above farmers' choice process is sufficiently stable across time for its main feature to be captured by parameters which are constant across time. This assumption set can hold for short run choices during a relatively small time period. Farms and farmers' production technology generally evolve slowly over time. This allows assuming that the parameters $\mathbf{q}_i$ of the production choice model remain constant over a few years. Short run production choices are repeated each year and follow the same scheme as long as the production technology and the quasi-fixed factor endowment do not change.

The second part of a parametric random parameter model describes the distribution of the farmers' specific parameters $\mathbf{q}_i$ conditional on the observed variables $\mathbf{z}_{it}$. It is assumed here that $\mathbf{z}_{it}$ and $\mathbf{q}_i$ are independent. This independence assumption holds, either if the heterogeneity control variable $\mathbf{c}_i$ contains the factors underlying the statistical dependence of $\mathbf{z}_{it}$ and $\mathbf{q}_i$, or if $\mathbf{z}_{it}$ does not vary across *i*. [2] It implies that one just needs a statistical model for $\mathcal{D}(\mathbf{q}_i)$. As in the empirical application, we define a parametric model for $\mathcal{D}(\mathbf{q}_i)$. This model is characterized by the density $h(\mathbf{q}_i;\boldsymbol{\eta})$ defined up to the parameter vector $\boldsymbol{\eta}$. $\mathcal{D}(\mathbf{q}_i)$ describes the distribution of $\mathbf{q}_i$ across the considered famers' population. The more the $\mathbf{q}_i$ varies across farmers, the more heterogeneity matters to model farmers' choices $\mathbf{c}_{it}$.

Of course the choice of the distribution $\mathcal{D}(\mathbf{q}_i)$ is crucial to suitably capture the unobserved heterogeneity effects in the considered model. Being related to unobserved variables, this choice basically is an empirical issue. It can be based on trials with different parametric models. Using flexible parametric models, *e.g.* finite discrete mixtures of Gaussian models, or non parametric models appears to be difficult in practice. Such models can only be used when the dimension of $\mathbf{q}_i$ is very small and with very large samples.

[2] In the empirical application presented in the next sections, $\mathbf{z}_{it}$ contains price variables which mostly vary across years and year dummies, ensuring that $\mathbf{q}_i$ and $\mathbf{z}_{it}$ can be considered as independent variables.

Specification of the role of $\mathbf{q}_i$ in the model of $\mathbf{c}_{it}$ depends on how unobserved heterogeneity effects are expected to affect farmers' choices. Standard panel data models generally assume that the effects of $\mathbf{q}_i$ and of $\mathbf{e}_{it}$ are additively separable in **r**, with *e.g.* $\mathbf{r}(\mathbf{z}_{it},\mathbf{e}_{it};\mathbf{q}_i)=\boldsymbol{\rho}(\mathbf{z}_{it})+\mathbf{q}_i+\mathbf{e}_{it}$. In this case the so-called "individual effect" $\mathbf{q}_i$ does not affect the effect of $\mathbf{z}_{it}$ on $\mathbf{c}_{it}$, implying homogeneous responses of $\mathbf{c}_{it}$ to changes in $\mathbf{z}_{it}$.
Keane (2009) discusses this point and highlights a basic trade-off. Econometric models defined as the sum of a deterministic part $\boldsymbol{\rho}(\mathbf{z}_{it})$ and of random terms $\mathbf{q}_i+\mathbf{e}_{it}$ are relatively easily estimated by using semi-parametric estimators. But such models do not suitably account for the effects of unobserved heterogeneity when these effects are not additively separable in the considered response functions, *i.e.* when $\frac{\partial}{\partial \mathbf{z}}\mathbf{r}(\mathbf{z}_{it},\mathbf{e}_{it};\mathbf{q}_i)$ actually depends on $\mathbf{q}_i$. Keane (2009) basically argues that the use of relatively involved inference tools as well as parametric assumptions on the distribution of the random terms $(\mathbf{e}_{it},\mathbf{q}_i)$ may be a reasonable price for buying the opportunity to introduce rich unobserved heterogeneity effects in the considered model.
Of course, this trade-off is an empirical issue and depends on the modeled choice process. But empirical evidences accumulated in other applied economics fields suggest that it is worth investigating this trade-off for agricultural production choice modeling. This is the main object of this article with a particular focus on the unobserved heterogeneity effects on farmers' responses to economic incentives.
The distribution of the dependent variable $\mathbf{c}_i$ conditional on its observed determinants $\mathbf{z}_i$, *i.e.* of $\mathcal{D}(\mathbf{c}_i\,|\,\mathbf{z}_i)$, is a key concept for estimation purposes. Its density defines the likelihood function to be used in the ML framework. The density of $\mathbf{c}_i$ conditional on $\mathbf{z}_i$ is the mean of that of $\mathbf{c}_i$ conditional on $(\mathbf{z}_i,\mathbf{q}_i)$ integrated over the distribution of $\mathbf{q}_i$:

(4) $$f(\mathbf{c}_i\,|\,\mathbf{z}_i;\boldsymbol{\theta})=\int f(\mathbf{c}_i\,|\,\mathbf{z}_i,\mathbf{q};\boldsymbol{\mu})h(\mathbf{q};\boldsymbol{\eta})d\mathbf{q}\,.$$

The term $\boldsymbol{\theta}\equiv(\boldsymbol{\mu}_1,...,\boldsymbol{\mu}_T,\boldsymbol{\eta})$ is the "complete" parameter vector of the considered parametric random parameter model. The integral in equation (4) cannot be solved analytically in general but this issue is ignored for the moment.

Statistical estimates of $\boldsymbol{\theta}$ allow the investigation of the distribution of the random parameter $\mathbf{q}_i$. First, these estimates can be used to test the empirical relevance of the random parameter specification by checking whether $\mathbf{q}_i$ exhibits true variations or not. If $\boldsymbol{\theta}$ contains "variance parameters" then simple parametric tests can be used. Second, the estimates of $\boldsymbol{\theta}$ can also be used to interpret the empirical content of the $\mathbf{q}_i$ terms. *E.g.*, the statistical relations among the elements of $\mathbf{q}_i$ may suggest interpretations of their variations.

The statistical estimates of $\boldsymbol{\theta}$ can also be used to "statistically calibrate" a simulation model based on the considered random parameter model. This simulation model can be based on estimates of the parameters for each sampled farmer. The distribution $\mathcal{D}(\mathbf{q}_i)$ is an *ex ante* or *prior* distribution of the random parameter. It describes the distribution of $\mathbf{q}_i$ in the considered farmer population. The minimum mean squared error estimator of $\mathbf{q}_i$ for any farmer taken at random in the considered population is simply the mean of $\mathbf{q}_i$, $E[\mathbf{q}_i]$. Of course, conditioning on the information set available for farmer *i* provides more accurate estimates, *i.e.* use of the information provided by $(\mathbf{c}_i,\mathbf{z}_i)$ allows defining more precise estimates of $\mathbf{q}_i$ based on $\mathcal{D}(\mathbf{q}_i \mid \mathbf{c}_i,\mathbf{z}_i)$.[3] By application of Bayes' rule the density of $\mathbf{q}_i$ conditional on $(\mathbf{c}_i,\mathbf{z}_i)$ is given by:

$$(5) \qquad h(\mathbf{q}_i \mid \mathbf{z}_i,\mathbf{c}_i;\boldsymbol{\theta}) = \omega(\mathbf{c}_i,\mathbf{z}_i,\mathbf{q}_i;\boldsymbol{\theta})h(\mathbf{q}_i;\boldsymbol{\eta}) \quad \text{where} \quad \omega(\mathbf{c}_i,\mathbf{z}_i,\mathbf{q}_i;\boldsymbol{\theta}) \equiv \frac{f(\mathbf{c}_i \mid \mathbf{z}_i,\mathbf{q}_i;\boldsymbol{\mu})}{f(\mathbf{c}_i \mid \mathbf{z}_i;\boldsymbol{\theta})}.$$

This density $h(\mathbf{q}_i \mid \mathbf{z}_i,\mathbf{c}_i;\boldsymbol{\theta})$ – which is designated as the *ex post* or *a posteriori* density of $\mathbf{q}_i$ (conditional on what is known about farmer *i*) – can be used to integrate $E[\mathbf{q}_i \mid \mathbf{c}_i,\mathbf{z}_i]$, the best predictor of $\mathbf{q}_i$ conditional on $(\mathbf{c}_i,\mathbf{z}_i)$ according to the minimum squared prediction error criterion[4].

We use the term "statistical calibration" instead of "estimation" to refer to the computation of the of $\mathbf{q}_i$ conditional on $(\mathbf{c}_i,\mathbf{z}_i)$. These predictions rely on $(\mathbf{c}_i,\mathbf{z}_i)$, the limited information available on farmer *i*, and are statistical in the sense that they depend on a

[3] Even a single observation $(\mathbf{c}_{it},\mathbf{z}_{it})$ for farmer *i* is valuable.

[4] Tell me what farm *i* has chosen, I'll tell you of which kind is likely to be farmer *i*. You can trust me because I know how farmers decide and because I know the characteristics of the farmers' population.

statistical estimate $\boldsymbol{\theta}$ as well as on the considered random parameter model which basically is a statistical model structured by a few micro-economic assumptions.

2. The random parameter multicrop model

The multicrop model considered here is a random parameter version of a model proposed by Carpentier and Letort (2013). This model combines a Nested MNL acreage share model with quadratic yield functions. This section presents the main features of the model to be used in the empirical application. Additional details related to this model and its theoretical background can be found in Carpentier and Letort (2013).

The considered multicrop model assumes that farmers maximize their expected profit in two steps. First they maximize the expected return to each crop under the assumption that this return doesn't depend on the crop acreages. Second, farmers allocate land to different crops to maximize their expected profit provided that they incur implicit acreage management costs. These management costs provide incentive for crop diversification.

The crop set $\mathcal{K} \equiv \{0,1,...,K\}$ is partitioned into mutually exclusive crop groups $\mathcal{K}_g$ for $g \in \mathcal{G} \equiv \{0,1,...,G\}$.[5] This partition is defined to account for the fact that different crops require different management efforts and compete more or less for, *e.g.*, quasi-fixed input uses. The groups are defined so that any crop compete more in the land allocation process with the other crops of its group than it does compete with crops of other groups. Group 0 contains a single crop, crop 0. As shown below, crop 0 plays a specific technical role in the model.

The "behavioral" model of the considered multicrop model is an equation system composed of a yield supply sub-system:

(6a) $\left\{ y_{k,it} = \beta_{k,i} + \delta_{k,t} - 1/2 \times \gamma_k w_{it}^2 p_{k,it}^{-2} + v_{k,it} \right.$ for $k \in \mathcal{K}_g$ and $g = 0,1,...,G$

and of an acreage share sub-system:

[5] This partition allows defining a two-level Nested acreage choice model. Further partitioning the crop groups allows defining multi-level Nested acreage choice model.

(6b)

$$\begin{cases} s_{k,it} = \dfrac{\exp(\rho_{g,i}\bar{\pi}_{k,it})}{\sum_{\ell\in\mathcal{K}_g}\exp(\rho_{g,i}\bar{\pi}_{\ell,it})} \dfrac{\left(\sum_{\ell\in\mathcal{K}_g}\exp(\rho_{g,i}\bar{\pi}_{\ell,it})\right)^{\alpha_i\rho_{g,i}^{-1}}}{\sum_{h\in\mathcal{G}}\left(\sum_{\ell\in\mathcal{K}_h}\exp(\rho_{h,i}\bar{\pi}_{\ell,it})\right)^{\alpha_i\rho_{h,i}^{-1}}} \text{ for } k\in\mathcal{K}_g \text{ and } g=1,\dots,G \\ \text{where } \bar{\pi}_{\ell,it} = p_{\ell,it}\left(\beta_{\ell,i}+\bar{\delta}_\ell\right)+1/2\times\gamma_\ell w_{it}^2 p_{\ell,it}^{-1} - \zeta_{\ell,i} + u_{\ell,it} \text{ for } \ell\in\mathcal{K}_g \text{ and } g=0,1,\dots,G \end{cases}$$

with $s_{0,it} = 1-\sum_{k=1}^{K} s_{k,it}$, by the total land use constraint.

Equation (6a) defines the yield supply of crop k as a function of the (anticipated) price of crop k, $p_{k,it}$, of the price of an aggregate variable input , w_{it}, and of an error term.

The yield function of crop k is obtained by maximizing in the aggregate variable input level, $x_{k,it}$, the expected margin of crop k, $\bar{\pi}_{k,it}$, under the assumptions that the yield function is quadratic in the aggregate variable input level:

(7) $$\left\{ y_{k,it} = \beta_{k,i} + \delta_{k,t} + v_{k,it} - 1/2\times\gamma_k^{-1}(\lambda_{k,i} + \upsilon_{k,it} - x_{k,it})^2 \right. \text{ with } E[v_{k,it}] = E[\upsilon_{k,it}] = 0$$

and that the random terms $\upsilon_{k,it}$ (which may include a year specific effect) are observed when the input level is decided.[6] This yield function is parameterized by two fixed parameters, the curvature parameter γ_k and the year specific yield effect $\delta_{k,t}$, and a random parameter $\beta_{k,i}$. It depends on the effects of random events represented by $v_{k,it}$ and $\upsilon_{k,it}$. Provided that γ_k needs to be positive for the yield function to be strictly concave, the farmer specific parameter $\beta_{k,i}$ can be interpreted as the maximum expected yield of crop k on farm i. This term depends on the natural endowment of the farm, on the production technology used by the farmer as well as on his ability.

The optimal input level of farmer i in t on crop k is thus given by:

(8) $$x_{k,it} = \lambda_{k,i} - \gamma_{k,0} w_{it} p_{k,it}^{-1} + \upsilon_{k,it}$$

and the corresponding expected gross margin is given by:

---

[6] Whether the random event effects $v_{k,it}$ and/or the year specific effects $\delta_{k,t}$ are observed or not doesn't matter. These effects are forgone by the considered (risk neutral) farmer.

(9) $\bar{\pi}_{k,it} = p_{k,it}\left(\beta_{k,i} + \overline{\delta_k}\right) + 1/2 \times \gamma_{k,0} w_{it}^2 p_{k,it}^{-1} - w_{it}\lambda_{k,i}$.

This gross margin level is expected by the farmer at the time of his acreage choices, *i.e.* before the observation of the random events represented by $v_{k,it}$ and $\upsilon_{k,it}$, and of the year specific effect $\delta_{k,t}$. The expectation across *t* of the year effect $\delta_{k,t}$ is set to be equal to an average year effect: $\bar{\delta}_k = T^{-1}\sum_t \delta_{k,t}$. Input demand equation (9) is not included in the estimated multicrop models because the input use levels are not observed at the crop level in our data set, they are only recorded at the farm level.[7]

Equation (6b) defines the acreage share optimal choices based on the expected profit maximization problem given by:

(10) $\max_{(s_k : k \in \mathcal{K}) \geq \mathbf{0}} \left\{ \sum_{k \in \mathcal{K}} s_k \bar{\pi}_{k,it} - C_i(s_k : k \in \mathcal{K}) \text{ s.t. } \sum_{k \in \mathcal{K}} s_k = 1 \right\}$.

The optimal acreage shares maximize the expected gross revenue of the farm, $\sum_{k \in \mathcal{K}} s_{k,it}\bar{\pi}_{k,it}$, minus the implicit management costs of the acreage choice, $C_i(s_{k,it} : k \in \mathcal{K})$ which is assumed to be strictly convex in the acreage share vector, under the total land use constraint, $\sum_{k \in \mathcal{K}} s_{k,it} = 1$. This constraint defines the crop 0 acreage share as a function of the other acreage shares with $s_{0,it} = 1 - \sum_{k=1}^{K} s_{k,it}$. The implicit management cost function C_i plays a crucial role here. Under the assumption that it is strictly convex in the acreage share vector $(s_k : k \in \mathcal{K})$, it formally defines the diversification motive of the crop acreage. It is defined by Carpentier and Letort (2013) as the sum of the unobserved costs and the shadow costs related to binding constraints due to limiting quasi-fixed factor quantities or to bio-physical factors.[8] Since the quasi-fixed factor endowments are highly heterogeneous, this cost function needs to be specified as farmer specific as much as possible. The functional form of the acreage share models in equation system (6a) are obtained by choosing the following Nested MNL management cost function:

[7] This aggregation problem can be overcome by defining an input use allocation equation as in Carpentier and Letort (2012). However, this option would have increased significantly the complexity of the considered multicrop model and of its estimation.

[8] It can also be interpreted as a penalty function for deviations from some reference acreage vector for which the quasi-fixed factor endowment of the farm is best suited.

$$C_i(s_k : k \in \mathcal{K}) = \sum_{g\in\mathcal{G}} \sum_{k\in\mathcal{K}_g} s_k(\chi_{k,i} + u_{k,it}) \tag{11}$$
$$+\alpha_i^{-1} \sum_{g\in\mathcal{G}} (1-\alpha_i\rho_{g,i}^{-1})\overline{s}_g \ln \overline{s}_g + \sum_{g\in\mathcal{G}} \rho_{g,i}^{-1} \sum_{k\in\mathcal{K}_g} s_k \ln s_k$$

up to an additive fixed cost. The term $\overline{s}_g \equiv \sum_{k\in\mathcal{K}_g} s_k$ defines the acreage share of group *g*. The strict convexity of C_i is ensured if $\rho_{g,i} \geq \alpha_i > 0$ for $g \in \mathcal{G}$. All parameters of C_i are assumed to be farmer specific to ensure its ability to capture the heterogeneity of the farms' capital endowments. The heterogeneity of the α_i and $\rho_{g,i}$ parameters plays a crucial role in this respect.[9] As shown by equation (6b), these terms largely determine the acreage choice elasticities. The larger they are, the more acreage choices are responsive to economic incentives. The $\chi_{k,i}$ parameters represent short run fixed costs.

Note also that the interpretation of C_i given above relies on the theoretical background given in Carpentier and Letort (2013). In empirical applications, the parameters of this function may also capture the effects of other diversification motives of crop acreages. *E.g.*, it may partly capture the effects of risk spreading motives (Chavas and Holt, 1990) or of crop rotations (Howitt, 1995). This provides further arguments for its specification based on farm specific parameters. In particular, farmers may have heterogeneous attitudes toward risk, financial constraints or personal wealth levels. This basically implies that the empirical estimates of the C_i functions need to be as reduced form functions capturing various diversification acreage motives while accounting for the heterogeneity of the effects of these motives among the considered farmers' population.

Many multicrop models proposed in the literature may use more flexible functional forms than the one considered here (see, *e.g.*, Chambers and Just, 1989 ; Oude Lansink and Peerlings, 1996 ; Moro and Skockai, 2006). As far as short run micro-economic choices are concerned, our viewpoint is that it may be more important to account for heterogeneity in the considered model than to use a highly flexible functional form for this model. Roughly speaking, if heterogeneity really matters it may be preferable to use a first order

[9] The term $\rho_{g,i}$ equals α_i if group *g* is a singleton. *E.g.*, we have $\rho_{0,i} = \alpha_i$. If $\rho_{0,i} = \alpha_i$ for $g \in \mathcal{G}$ then equation (6b) gives the Standard MNL acreage share model and equation (11) gives the corresponding acreage management cost function.

approximation for each sampled farm rather than to use a second order approximation defined at the sample level.

Due to insufficient variation of the aggregate input prices in our data set,[10] it is difficult to separately identify the $\lambda_{k,i}$ and $\chi_{k,i}$ parameters empirically. This explains why the expected gross margin used in the acreage share model (6b) is not that given by equation (9). The term $\zeta_{k,i}$ in equation (6b) is given by $\zeta_{k,i} \simeq \chi_{k,i} - w_{it}\lambda_{k,i}$.

Note also that the total land use constraints imply that the terms $\zeta_{k,i}$ and $u_{k,it}$ are defined up to an additive term. The terms $\zeta_{0,i}$ and $u_{0,it}$ are imposed to be null to overcome this identification problem.[11]

Additional notations are required to present the distributional assumptions defining the parametric model considered in the empirical application. The following system level vectors are obtained from the crop level parameters: $\mathbf{s_{it}} \equiv \left(s_{k,it} : k \in \mathcal{K}\right)$, $\mathbf{y}_{it} \equiv (y_{k,it} : k \in \mathcal{K})$, $\mathbf{p}_{it} \equiv (p_{k,it} : k \in \mathcal{K})$, $\mathbf{v}_{it} \equiv (v_{k,it} : k \in \mathcal{K})$, $\mathbf{u}_{it} \equiv (u_{k,it} : k \in \mathcal{K} \setminus \{0\})$, $\boldsymbol{\beta}_i \equiv (\beta_{k,i} : k \in \mathcal{K})$, $\boldsymbol{\rho}_i \equiv (\rho_{g,i} : g \in \mathcal{G} \setminus \{0\})$, $\boldsymbol{\zeta}_i \equiv (\zeta_{k,i} : k \in \mathcal{K} \setminus \{0\})$, $\boldsymbol{\delta}_t \equiv (\delta_{k,t} : k \in \mathcal{K})$, and $\boldsymbol{\gamma} \equiv (\gamma_k : k \in \mathcal{K})$. The vector $\boldsymbol{\delta} \equiv (\boldsymbol{\delta}_t : t = 2,...,T)$ contains the year specific effects on the yield supply functions.

In order to relate the multicrop model in this section with the generic choice model in the preceding section, we finally define the farmer choice vector $\mathbf{c}_{it} \equiv (\mathbf{y}_{it}, \mathbf{s}_{it})$, exogenous variable vector $\mathbf{z}_{it} \equiv (\mathbf{p}_{it}, w_{it}, d_{it})$, error term vector $\mathbf{e}_{it} \equiv (\mathbf{v}_{it}, \mathbf{u}_{it})$ and random parameter vector $\mathbf{q}_i \equiv (\ln \boldsymbol{\beta}_i, \ln \alpha_i, \ln \boldsymbol{\rho}_i, \boldsymbol{\zeta}_i)$. The exogenous variable vector $\mathbf{z}_{it}$ includes the year dummy variable d_t.

The response function **r** considered in the preceding section is given by equations (6). It is parameterized by the random parameter vector $\mathbf{q}_i$. The counterpart of **r** in the multicrop model is also parameterized by the fixed parameter vector $\tau \equiv (\boldsymbol{\gamma}, \boldsymbol{\delta})$. As argued in the preceding section, the terms $\mathbf{z}_{it}$, $\mathbf{e}_{it}$ and $\mathbf{q}_i$ are assumed to be mutually independent for $t = 1,...,T$. Provided that the stochastic events affecting the crop production process are

---

[10] As well as due to our not modeling input demands.

[11] These normalization constraints imply that the $\zeta_{k,i}$ and $u_{k,it}$ terms are to be interpreted as differences with their counterparts for crop 0.

unknown at the time acreage choices are made, it can be assumed that the terms $\mathbf{v}_{it}$ and $\mathbf{u}_{it}$ are also independent for $t=1,...,T$.

As is standard for error terms, $\mathbf{v}_{it}$ and $\mathbf{u}_{it}$ are assumed to be normal with $\mathbf{v}_{it} \sim \mathcal{N}(\mathbf{0},\mathbf{\Lambda})$ and $\mathbf{u}_{it} \sim \mathcal{N}(\mathbf{0},\mathbf{\Psi})$. The mixing distribution of the model is also assumed to be normal with $\mathbf{q}_i \sim \mathcal{N}(\mathbf{a},\mathbf{\Omega})$. The covariance matrix $\mathbf{\Omega}$ being unrestricted this probability distribution imposes no restriction on the relationships among the elements of $\mathbf{q}_i$. Due to the log transformation of $\boldsymbol{\beta}_i$, α_i and $\boldsymbol{\rho}_i$ in $\mathbf{q}_i$, these terms are indeed assumed to be jointly log-normal. This ensures their strict positivity.

Once again, in order to related the multicrop model considered here to the more general framework elements of the previous section, the distinct elements of the parameters $\boldsymbol{\alpha}$, $\mathbf{\Lambda}$ and $\mathbf{\Psi}$ are collected in $\boldsymbol{\mu}$, those of $\mathbf{a}$ and $\mathbf{\Omega}$ are collected in $\boldsymbol{\eta}$, and $\boldsymbol{\theta} \equiv (\boldsymbol{\mu},\boldsymbol{\eta})$ defines the "full" parameter vector to be estimated.

The inverse function of $\mathbf{r}$ is required to determine the (conditional) likelihood functions of the considered model. The elements of $\mathbf{v}_{it}$ can easily be recovered with:

(12a) $\quad v_{k,it} = y_{k,it} - \beta_{k,i} - \delta_{k,t} + 1/2 \times \gamma_k w_{it}^2 p_{k,it}^{-2}$

while the elements of $\mathbf{u}_{it}$ can be obtained by application of Berry's (1994) device:

(12b) $\quad u_{k,it} = \begin{bmatrix} \dfrac{1}{\alpha_i}\left(\ln s_{k,it} - \ln s_{0,it} - \left(1-\alpha_i\rho_{g,i}^{-1}\right)\times\left(\ln s_{k,it} - \ln \overline{s}_{g,it}\right)\right) + \zeta_{k,i} \\ -p_{k,it}\left(\beta_{k,it}+\overline{\delta}_k\right) - 1/2\times\gamma_k w_{it}^2 p_{k,it}^{-1} + p_{0,it}\left(\beta_{0,it}+\overline{\delta}_0\right) + 1/2\times\gamma_0 w_{it}^2 p_{0,it}^{-1} \end{bmatrix}.$

The density of $\mathbf{c}_i$ conditional on $(\mathbf{z}_i,\mathbf{q}_i)$ can be obtained by applying equations (2) and (3) and by using the density of normal random vectors. Let $\varphi(\mathbf{u};\mathbf{B})$ denote the density function of $\mathcal{N}(\mathbf{0},\mathbf{B})$ at $\mathbf{u}$. The density of $\mathbf{y}_{it}$ conditional on $(\mathbf{z}_i,\mathbf{q}_i)$ is given by:

(13a) $\quad f(\mathbf{y}_{it} \mid \mathbf{z}_i,\mathbf{q}_i;\boldsymbol{\mu}) = \varphi(\mathbf{v}_{it};\mathbf{\Lambda})$

and that of $\mathbf{s}_{it}$ conditional on $(\mathbf{z}_i,\mathbf{q}_i)$ is given by:

(13b) $\quad f(\mathbf{s}_{it} \mid \mathbf{z}_i,\mathbf{q}_i;\boldsymbol{\mu}) = \alpha_i^{G-1}\left(\prod_{g\in\mathcal{G}} \rho_{g,i}^{K_g-1}\right)\left(\prod_{k\in\mathcal{K}} s_{k,it}^{-1}\right)\times\varphi(\mathbf{u}_{it};\mathbf{\Psi})$.

We obtain that:

(13c) $\quad f(\mathbf{c}_{it} \mid \mathbf{z}_i,\mathbf{q}_i;\boldsymbol{\mu}) = f(\mathbf{y}_{it} \mid \mathbf{z}_i,\mathbf{q}_i;\boldsymbol{\mu})\, f(\mathbf{s}_{it} \mid \mathbf{z}_i,\mathbf{q}_i;\boldsymbol{\mu})$

thanks to the mutual independence of $\mathbf{v}_{it}$ and $\mathbf{u}_{it}$ conditional on $(\mathbf{z}_i, \mathbf{q}_i)$ for $t = 1,...,T$. Finally, the random parameter vector density is given by:

(13d) $h(\mathbf{q}_i;\boldsymbol{\eta}) = \varphi(\mathbf{q}_i - \mathbf{a};\boldsymbol{\Omega})$.

The basis of the estimator of the interest parameter is the likelihood function of the farmers' choice sequence $\mathbf{c}_i$ conditional on $\mathbf{z}_i$ at $\boldsymbol{\theta}$ and is equal to $f(\mathbf{c}_i \mid \mathbf{z}_i;\boldsymbol{\theta})$. This function can be obtained by considering equations (13) and (4).

3. Estimation issues

Estimation of a random parameter model such as the one presented in the preceding section requires specific estimators due to its specific structure. From a theoretical viewpoint, the parameters of this fully parametric model can be efficiently estimated according to the ML principle. But the ML estimator of $\boldsymbol{\theta}$ is practically "infeasible" because the individual likelihood functions, *i.e.* the $f(\mathbf{c}_i \mid \mathbf{z}_i;\boldsymbol{\theta}) = \int f(\mathbf{c}_i \mid \mathbf{z}_i, \mathbf{q};\boldsymbol{\mu})h(\mathbf{q};\boldsymbol{\eta})d\mathbf{q}$ terms, cannot be integrated in our case, neither analytically, nor numerically. These likelihood functions must be integrated with simulation methods, implying that the estimators of $\boldsymbol{\theta}$ must be simulated counterparts of the standard ML estimators. Furthermore, maximizing the "true" sample log-likelihood function, *i.e.* $\sum_{i=1}^{N} \ln f(\mathbf{c}_i \mid \mathbf{z}_i;\boldsymbol{\theta})$, in $\boldsymbol{\theta}$ would be very difficult in practice, due to the functional form of the individual likelihood functions and due to the dimension of $\boldsymbol{\theta}$. Statisticians have proposed specific extensions of the Expectation-Maximization (EM) algorithm of Dempster *et al* (1977) to compute the ML estimators of random parameter (or mixed) models. We employ a Simulated EM (SEM) algorithm to compute an estimator whose asymptotic properties are basically those of the infeasible ML estimator of $\boldsymbol{\theta}$ (see, *e.g.*, McLachlan and Krishnan (2008) for a recent review of the numerous SEM algorithms proposed in the statistics literature).

This section presents the main features of our computation strategy for our estimator of $\boldsymbol{\theta}$. The choice of this computation strategy, *i.e.* the particular design of the SEM algorithm we use, was mainly based on practical arguments. *E.g.* other SEM algorithms may be more efficient from a numerical viewpoint or may require less simulations, and thus less computing time or power, to perform well. But this algorithm is relatively easy to code, has good theoretical properties and seems to perform well in practice, at least as far as our limited experience proves this. Further details are available from the authors upon request.

Our estimators are built by estimating $f(\mathbf{c}_i \mid \mathbf{z}_i;\boldsymbol{\theta})$ with simulation methods. Provided that the $\tilde{\mathbf{q}}_{i,s}(\boldsymbol{\eta})$ terms are independent random draws from $h(\mathbf{q};\boldsymbol{\eta})$ for $s=1,...,S$, the strong law of large numbers guarantees that:

(14) $$\tilde{f}_S(\mathbf{c}_i \mid \mathbf{z}_i;\boldsymbol{\theta}) \equiv S^{-1}\sum\nolimits_{s=1}^{S} f(\mathbf{c}_i \mid \mathbf{z}_i, \tilde{\mathbf{q}}_{i,s}(\boldsymbol{\eta});\boldsymbol{\mu})$$

almost surely converges to $f(\mathbf{c}_i \mid \mathbf{z}_i;\boldsymbol{\theta})$ as S rises to infinity. While econometricians usually employ Simulated ML (SML) estimators in this context, statisticians usually prefer to rely on SEM algorithms to compute estimators which differ from SML estimators but which basically share the same asymptotic properties as S and N grows to infinity, with S rising faster than $\sqrt{N}$ (Jank and Booth, 2003). The SML estimator of $\boldsymbol{\theta}$ is obtained by directly maximizing the sample simulated log-likelihood function $\ln \tilde{L}_{S,N}(\boldsymbol{\theta}) \equiv \sum\nolimits_{i=1}^{N} \ln \tilde{\ell}_S(\boldsymbol{\theta};\mathbf{c}_i \mid \mathbf{z}_i)$, usually by relying on gradient-based algorithms. This maximization problem is difficult to solve in our empirical application because $\ln \tilde{L}_{S,N}(\boldsymbol{\theta})$ is highly non linear in $\boldsymbol{\theta}$ and because the dimension of $\boldsymbol{\theta}$ is quite large.[12]

The EM algorithm is particularly well suited to compute ML estimators in cases where the model of interest involves hidden variables such as random parameters. It consists in iterating two steps, the Expectation step (E step) and the Maximization step (M step), until numerical convergence. It basically replaces a large ML problem by a sequence of simpler maximization problems. [13]

In our case the EM algorithm involves the following density:

(15) $$\kappa(\mathbf{c}_i, \mathbf{q}_i \mid \mathbf{z}_i;\boldsymbol{\theta}) \equiv f(\mathbf{c}_i \mid \mathbf{z}_i, \mathbf{q}_i;\boldsymbol{\mu})h(\mathbf{q}_i;\boldsymbol{\eta}).$$

[12] *E.g.*, Train (2009) reports that the variance matrix of Gaussian mixing probability distribution is not easily recovered by SML estimators, leading to the restriction that this matrix is diagonal or block-diagonal in many empirical studies.

[13] The EM algorithm also increases the sample log-likelihood at each iteration, implying that it generally leads to a (local) maximum of the considered likelihood function. SEM algorithms do not necessarily monotonically increase the simulated sample log-likelihood due to the simulation noise. The main drawback of the EM algorithm is that, albeit it moves quickly into the neighborhood of ML estimator of $\boldsymbol{\theta}$, it numerically converges slowly within this neighborhood.

The term $\kappa(\mathbf{c}_i, \mathbf{q}_i \mid \mathbf{z}_i; \boldsymbol{\theta})$ is the distribution function of the "complete" dependent variable vector $(\mathbf{c}_i, \mathbf{q}_i)$ conditional on the exogenous variable $\mathbf{z}_i$. As a result $\ln \kappa(\mathbf{c}_i, \mathbf{q}_i \mid \mathbf{z}_i; \boldsymbol{\theta})$, is the log-likelihood function at $\boldsymbol{\theta}$ of $(\mathbf{c}_i, \mathbf{q}_i)$ conditional on $\mathbf{z}_i$. At iteration *n,* provided that $\boldsymbol{\theta}_{n-1}$ is the value of $\boldsymbol{\theta}$ obtained at the end of iteration $n-1$, the EM algorithm iterates the following steps until numerical convergence:

E step$_n$. Integration of the conditional expectations:

(16) $E[\ln \kappa(\mathbf{c}_i, \mathbf{q}_i \mid \mathbf{z}_i; \boldsymbol{\theta}) \mid \mathbf{z}_i, \mathbf{c}_i; \boldsymbol{\theta}_{n-1}] \equiv \int \ln \kappa(\mathbf{c}_i, \mathbf{q} \mid \mathbf{z}_i; \boldsymbol{\theta}) h(\mathbf{q} \mid \mathbf{z}_i, \mathbf{c}_i; \boldsymbol{\theta}_{n-1}) d\mathbf{q}$ for $i = 1, \ldots, N$.

M step$_n$. Update of the value of $\boldsymbol{\theta}$ with:

(17) $\boldsymbol{\theta}_n \equiv \arg\max_{\boldsymbol{\theta}} Q_N(\boldsymbol{\theta} \mid \boldsymbol{\theta}_{n-1})$ where $Q_N(\boldsymbol{\theta} \mid \boldsymbol{\theta}_{n-1}) \equiv \sum_{i=1}^{N} E[\ln \kappa(\mathbf{c}_i, \mathbf{q}_i \mid \mathbf{z}_i; \boldsymbol{\theta}) \mid \mathbf{z}_i, \mathbf{c}_i; \boldsymbol{\theta}_{n-1}]$.

The E step thus consists in integrating the individual log-likelihood functions $\ln \kappa(\mathbf{c}_i, \mathbf{q}_i \mid \mathbf{z}_i; \boldsymbol{\theta})$ over the *ex post* density $h(\mathbf{q}_i \mid \mathbf{z}_i, \mathbf{c}_i; \boldsymbol{\theta}_{n-1})$. This integration yields the expectation of log-likelihood function at $\boldsymbol{\theta}$ of the "complete" dependent variable vector of farmer *i* conditional on what is known on this farmer, *i.e.* $(\mathbf{z}_i, \mathbf{c}_i)$, and assuming that $\boldsymbol{\theta}_{n-1}$ is the true value of the interest parameter. The updated value of $\boldsymbol{\theta}_n$ is then defined as an ML estimator of based on the individual expected log-likelihood functions computed in the E step.

Equation (15) is specific to models involving hidden variables. It is used to split the M step into two maximization problems:

(18a) $\boldsymbol{\mu}_n \equiv \arg\max_{\boldsymbol{\mu}} \sum_{i=1}^{N} E[\ln f(\mathbf{c}_i \mid \mathbf{z}_i, \mathbf{q}_i; \boldsymbol{\mu}) \mid \mathbf{z}_i, \mathbf{c}_i; \boldsymbol{\theta}_{n-1}]$

and:

(18b) $\boldsymbol{\eta}_n \equiv \arg\max_{\boldsymbol{\eta}} \sum_{i=1}^{N} E[\ln h(\mathbf{q}_i; \boldsymbol{\eta}) \mid \mathbf{z}_i, \mathbf{c}_i; \boldsymbol{\theta}_{n-1}]$

where $\boldsymbol{\theta}_n \equiv (\boldsymbol{\mu}_n, \boldsymbol{\eta}_n)$. Basically, the parameters of the "behavioral model" on the one hand, and those of the "mixing" model on the other hand can be separately updated. In our case, the elements of $\boldsymbol{\eta}_n$ are defined as empirical means and covariances.

Of course, the expectations in equations (16)–(18) cannot be computed neither analytically, nor numerically. The EM algorithm described above would lead to the infeasible ML estimator of $\boldsymbol{\theta}_n$. The SEM algorithms were proposed to extend the use of the EM algorithms in cases where the Expectation step requires integration by simulation methods.[14]

Our estimates were computed by using an algorithm in the class of the SEM algorithms proposed by Delyon *et al* (1999). These algorithms, designated as the Stochastic Approximation EM (SAEM) algorithms, have two main advantages. First, they are numerically stable despite their relying on integration by simulation methods at each of their iterations. Second, they allow using simplified versions of the M step. The M step presented in equations (17) defines $\boldsymbol{\theta}_n$ as the maximand in $\boldsymbol{\theta}$ of $Q_N(\boldsymbol{\theta}\,|\,\boldsymbol{\theta}_{n-1})$. Indeed, $\boldsymbol{\theta}_n$ can just be defined as a value of $\boldsymbol{\theta}$ such that $Q_N(\boldsymbol{\theta}_n\,|\,\boldsymbol{\theta}_{n-1}) > Q_N(\boldsymbol{\theta}_{n-1}\,|\,\boldsymbol{\theta}_{n-1})$, *i.e.* such that $\boldsymbol{\theta}_n$ simply increases the value of $Q_N(\boldsymbol{\theta}\,|\,\boldsymbol{\theta}_{n-1})$ from $Q_N(\boldsymbol{\theta}_{n-1}\,|\,\boldsymbol{\theta}_{n-1})$.[15] We used the simplification of the M step proposed by Meng and Rubin (1993), *i.e.* the sequence of Conditional Maximization (CM) steps of their Expectation–Conditional Maximization (ECM) algorithm. The proposed algorithm only involves simple arithmetic operations, *i.e.* the ones required to compute empirical means and OLS estimators. Finally, the expectations in equation (18) were integrated by using the simulator proposed by Train (2007, 2008).

Our detailed estimation procedure is available upon request. We briefly present Train's simulator because it was used to estimate the conditional expectations in equations (16)–(18) as well as to calibrate the farmer's specific parameters in our empirical application, the complete algorithm used to compute the estimates of the empirical application being available upon request.

Train's simulator allows estimating the expectation of any function τ of $(\mathbf{q}_i, \mathbf{z}_i, \mathbf{c}_i)$, $\tau_i(\mathbf{q}_i) \equiv \tau(\mathbf{q}_i, \mathbf{z}_i, \mathbf{c}_i)$, integrated over the *ex post* density of the random parameters

[14] Note that equations (4) and (5) show that the integration problems encountered either when using the EM algorithm or when considering direct ML procedures have the same root, *i.e.* it is difficult to compute $h(\mathbf{q}_i\,|\,\mathbf{z}_i, \mathbf{c}_i; \boldsymbol{\theta})$ because it is difficult to compute $f(\mathbf{c}_i\,|\,\mathbf{z}_i; \boldsymbol{\theta})$.

[15] In their seminal article, Dempster *et al* (1997) also considered this extension of the standard M step to define an extension of the standard EM algorithm which they designated as the Generalized EM (GEM) algorithm.

$h(\mathbf{q}_i \mid \mathbf{z}_i, \mathbf{c}_i; \boldsymbol{\theta})$ by simply using independent random draws from their *ex ante* density $h(\mathbf{q};\boldsymbol{\eta})$, *i.e.* by using the $\tilde{\mathbf{q}}_{i,s}(\boldsymbol{\eta})$ draws.[16] Equation (5) allows showing that:

(19) $$E[\tau_i(\mathbf{q}_i) \mid \mathbf{z}_i, \mathbf{c}_i; \boldsymbol{\theta}] \equiv \int \tau_i(\mathbf{q}) h(\mathbf{q} \mid \mathbf{z}_i, \mathbf{c}_i; \boldsymbol{\theta}) d\mathbf{q} = \int \omega_i(\boldsymbol{\theta}) \tau_i(\mathbf{q}_i) h(\mathbf{q};\boldsymbol{\eta}) d\mathbf{q},$$

The strong law of large numbers then ensures that:

(20a) $$\tilde{E}_S[\tau_i(\mathbf{q}_i) \mid \mathbf{z}_i, \mathbf{c}_i; \boldsymbol{\theta}] \equiv S^{-1} \sum_{s=1}^{S} \tilde{\omega}_{i,s}(\boldsymbol{\theta}) \tau_i(\tilde{\mathbf{q}}_{i,s}(\boldsymbol{\eta}))$$

where:

(20b) $$\tilde{\omega}_{i,s}(\boldsymbol{\theta}) \equiv \frac{f(\mathbf{c}_i \mid \mathbf{z}_i, \tilde{\mathbf{q}}_{i,s}(\boldsymbol{\eta}); \boldsymbol{\mu})}{S^{-1} \sum_{s=1}^{S} f(\mathbf{c}_i \mid \mathbf{z}_i, \tilde{\mathbf{q}}_{i,s}(\boldsymbol{\eta}); \boldsymbol{\mu})}$$

almost surely converges to $E[\tau_i(\mathbf{q}_i) \mid \mathbf{z}_i, \mathbf{c}_i; \boldsymbol{\theta}]$ as S rises to infinity. The use of the weight terms $\tilde{\omega}_{i,s}(\boldsymbol{\theta})$ show that Train's simulator can be interpreted as an importance sampling simulator with $h(\mathbf{q}_i;\boldsymbol{\eta})$ as the proposal density. This proposal distribution clearly is inefficient, *i.e.* if $\mathbf{q}_i$ exhibits significant variability then $h(\mathbf{q}_i;\boldsymbol{\eta})$ is unlikely to be close to $h(\mathbf{q}_i \mid \mathbf{z}_i, \mathbf{c}_i; \boldsymbol{\theta})$, but the simplicity of the proposed algorithm allows using very large random draw numbers for approximating $f(\mathbf{c}_i \mid \mathbf{z}_i; \boldsymbol{\theta})$.

4. Empirical application

As an illustrative application of the approach proposed in this paper to account for farm heterogeneity, we use a set of French data to estimate the multicrop model presented in the second section. These estimations allow an investigation of the distribution of the random parameters of the model, which comes to illustrate the importance of unobserved heterogeneity in farmers' production choices. Based on these estimation results, we perform a "statistical calibration" of the model parameters for each sampled farmer in order (i) to evaluate the performances of the estimated model and (ii) to reveal some

---

[16] Such expectations can be integrated by using draws from $h(\mathbf{q}_i \mid \mathbf{z}_i, \mathbf{c}_i; \boldsymbol{\theta})$ which are more difficult to obtain. *E.g.*, it is always possible to obtain Metropolis-Hastings random draws from $h(\mathbf{q}_i \mid \mathbf{z}_i, \mathbf{c}_i; \boldsymbol{\theta})$. But this simulation technique consists in a rather long process to be repeated at each iteration of the SEM algorithm. Train's simulator appears to be much more convenient. Random draws from $h(\mathbf{q}_i;\boldsymbol{\theta})$ are easily obtained with random draws from the standard uniform distribution. Furthermore, the same draws from $h(\mathbf{q}_i;\boldsymbol{\theta})$ can be used for each farmer for an iteration and/or the random draws from the standard uniform can be re-used to compute the draws from $h(\mathbf{q}_i;\boldsymbol{\theta})$ along the SEM algorithm.

potential determinants of the heterogeneity in farmers' behaviors. We then perform some simulations in order to study the impacts and potential implications of the modeling of heterogeneous behaviors on simulation results.

4.1. Data

The data set used to estimate our model is a panel data sample of 391 observations of French grain crop producers in the large (geological) Paris basin over the years 2004 to 2007, obtained from the Farm Accountancy Data Network (FADN). It provides detailed information on crop production for each farm: acreage, yield and price at the farm gate. The aggregated input price index is made available at the regional level by the French Department of Agriculture.

In our application yield levels and acreage share choices are considered for three (aggregated) crops: soft wheat (crop 1), other cereals (mainly barley and corn, crop 2) and, oilseeds (mainly rapeseed) and protein crops (mainly peas) (crop 0). Crop aggregates are based on agronomic considerations. The basic rotation scheme of the French grain producers is a sequence with three crops as: rootcrops (*e.g.* potato or sugar beet) or protein crop or oilseed (*e.g.* rapeseed or sunflower) – winter wheat – secondary cereal (*e.g.* barley or wheat). This scheme is adapted to soil and climatic conditions. Rootcrops require good quality soils which are found in the north of France. Sunflower is grown in the south of France while rapeseed, the other main oilseed crop is grown in the north of France (our region of interest). Sugar beet and potato acreages were considered exogenous due to production quotas for sugar beet and production contracts for potatoes.

The considered sample only includes observations with strictly positive acreages. This selection rule doesn't lead to significant attrition thanks to the crop aggregation procedure. Our sample covers the French regions specialized in grain production, with the notable exception of the south-west of France where corn monoculture is the dominant cropping system. Farms are observed for 3 years on average. We assume that farms' attrition is exogenous. The French FADN is constructed as a rotating panel seeking to collect data for 4 years for each sampled farm. Such an attrition is easily accommodated in our modeling framework. Farms' likelihood functions are computed according to the observed choice sequences.

4.2. Estimation Results

Our estimations are conducted by using the SAS software and applying the procedure presented in the third part of the paper. The recursive step of simulation of SAEM algorithm is implemented using 1000 draws. The algorithm converges without difficulties after 244 iterations. Results were not significantly affected by the use of alternative starting values or the use of larger number of draws.

Selected estimation results are reported in Table 1 and Table 2, the complete results being available from the authors upon request. These results show that the model fits relatively well to the data. Indeed, most parameters, especially the expectations and covariances of the random parameters and the variance matrices of the error terms appear to be precisely estimated. The fixed parameters representing price ($\boldsymbol{\gamma}$) and time ($\boldsymbol{\delta}$) effects appear to be less precisely estimated. This is due to a lack of variation in crop prices in the period covered by our data: price effects on yields can hardly be distinguished from time-related effects. Since climatic events are the main sources of yield variations from one year to another, this issue could be overcome by introducing climatic variables in the model. This is however not the central in the present study, which aims at exploring the heterogeneous determinants of farmers' behaviors. Indeed, even if price and climatic effects are not separately identified, our estimation account for their joint impact on yields; there is thus no reason for the introduction of climatic variables to change the results of the estimations of the random parameters distribution.

The yield equation parameters (reported in Table 1) are precisely estimated. This was expected since each yield equation basically is a regression equation with individual random terms. The parameter estimates lie in reasonable ranges. The estimates of the probability distribution of $\boldsymbol{\beta}_i$ show that the $\beta_{k,i}$ parameters significantly vary across farms while being strongly positively correlated to each other. This was expected because yield potentials vary across regions, and because good growing conditions for a grain crop are also good for the others. The variance of $\beta_{k,i}$ is higher or close to that of $v_{k,it}$ for wheat and other cereals, but the variance of $v_{k,it}$ is twice that of error terms in the oilseeds case. This may reflects at least two points: first, a large part of the heterogeneity in cereals, and notably wheat, yields is due differences in unobservable characteristics of each farm or farmer; second, provided that rapeseed is by far the most important oilseed in northern France, these results may be due the fact that the rapeseed yield is more risky than the cereal yield, mostly due to bugs and diseases.

The acreage share equation parameter estimates (reported in Table 2) also range in reasonable ranges. The estimated expectation of $\ln \alpha_i$, respectively of $\ln \rho_i$, equals -2.357, respectively -2.186. Importantly, the estimate of the expectation of ρ_i is higher than that of α_i. This is a sufficient condition for the entropic acreage management cost function, lying at the root of the Nested MNL acreage share function, to be convex. According to the estimates of their respective variances, the α_i and ρ_i parameters significantly vary across farms. This result is important for simulation studies because these parameters largely determine acreage price elasticities in MNL acreage share models. The higher α_i and ρ_i are, the more reactive the acreages are to price changes. The elements of $\boldsymbol{\beta}_i$ appear to be positively correlated with α_i. A possible interpretation of this result is as follows. High levels of $\boldsymbol{\beta}_i$ indicate good farming conditions for grain crops in farm i and/or farmer i technical ability. This implies that the farm operation is sufficiently profitable to allow suitable machinery investments which, in turn, implies a high level of α_i and, finally, relatively unconstrained acreage choices between cereals and oilseeds. The results are different when it comes to acreage adjustments within the cereal nest: the elements of $\boldsymbol{\beta}_i$ are not positively correlated with ρ_i, which tends to show that the flexibility of acreage adjustments between wheat and other cereals is associated to other factor than the one advocated previously.

4.3. Statistical calibration of individual parameters

As explained in the first part of the paper, the estimated parametric model allows a computation of the (random) $\mathbf{q}_i$ parameters for each farm/farmer of the sample, according to the logic “tell me what you do, I'll tell you who you are”. Once the *ex ante* distribution of $\mathbf{q}_i$ in the population has been estimated we “statistically calibrate” the specific parameters for each individual i based on the *ex post* density of $\mathbf{q}_i$, that is conditional on observed farmers' responses to economic incentives. The *ex post* and *ex ante* density of selected random parameters ($\boldsymbol{\beta}_i$, α_i and ρ_i) are represented on Figure 1. The two distributions almost superimposed for all parameters, which reflects a good specification of our model (Train, 2007). We can also notice that the distributions of the $\boldsymbol{\beta}_i$ parameters, representing maximum potential yields of the farms, appear to be more spread for other cereals than for the two other crops, reflecting a higher heterogeneity of yields between farms for that crop. That might be due to the fact that "other cereals" is an aggregate of various crops (mainly corn and barley), whereas "wheat" is a single crop and "oilseeds" is essentially composed of rapeseed in our sample. The α_i and ρ_i parameters exhibit right-skewed distribution. The two distributions reflects the fact that α_i parameters generally take lower values than ρ_i parameters (this is actually the case for 73% of the farms/farmers, the remaining 27% individuals having ρ_i values almost equal to α_i values), which reflects more flexible adjustments between wheat and other cereal acreages than between oilseeds and other crops and is a sufficient condition for the acreage management cost function to be convex.

Figure 2 reports the calibrated values of the $\boldsymbol{\beta}_i$, α_i and ρ_i parameters together with their confidence intervals for each farm/farmer of the sample. We can see from these graphs that confidence intervals of parameters do not overlap for all individuals: these parameters do actually take different values from one individual to another. This comes to illustrate the heterogeneity in potential yields across farms and in the way farmers are able to adjust their acreages in response to economic incentives.

Having calibrated individual parameters for each farm/farmer, we are able to compute the individual yields and acreages predicted by the NMNL model. Based on these predictions, we can then compute "pseudo R^2" criteria corresponding to the share of the variance of interest variables predicted by the model, and compare the average observed values of these variables to their predicted values. These fitting criteria of the model are reported in Table 3 below. Once again, the model proves to fit well the data, especially for wheat and other cereals with "pseudo R^2" around 60% for yields and 70% for acreage shares, and observed and predicted average values very close one to each other for all interest variables.

Up to this point, we have demonstrated that farmers' behaviors do actually rely on heterogeneous determinants which are not explicitly introduced in the model used to represent their production choices. It thus seems crucial to account for heterogeneity in micro econometric production choice models. If the sources of this heterogeneity were known to econometricians, they could be controlled for through, for instance, the use of control variables[17]. However, if some of them are identifiable, heterogeneity sources are multiple and most of them can certainly not be reduced to farm/farmers' observable characteristics. This point is illustrated by Figure 3 and Table 4.

Maps reported on Figure 3 show the calibrated values of three parameters: $\beta_{k,i}$ for wheat, α_i and ρ_i[18] for each farm of our sample. The top left map clearly shows that the distribution of potential wheat yields exhibits a spatial pattern, the highest yields being located in the North of France. This is in total accordance with what is known about the different agronomic potentials of French regions. Introducing spatial farm characteristics in the model could help accounting for some heterogeneity. Farms' localization is however not the only source of heterogeneity in agricultural production choices. This is reflected by the two other maps on Figure 3: the distribution of the α_i and ρ_i parameters across space is different from that of the $\beta_{k,i}$ parameters. No specific spatial pattern seems emerge from these maps.

In a further attempt to qualify the potential sources of farmers' behaviors heterogeneity, we have computed the correlation between the values of individual parameters and some observable farms/farmers characteristics considered as exogenous in the model: the amount of farm capital, the root crops acreage and the age of farmer[19]. Farm capital is positively and significantly correlated with the $\beta_{k,i}$ parameter for oilseed, the α_i parameter, and to a lesser extent the $\beta_{k,i}$ parameter for wheat. This reflects one argument previously advocated: farms endowed with more capital are the more productive ones and also own

[17] Of course the use of control variables is allowed in our modeling approach. But it is omitted for simplicity as well as for investigating the potential of random parameter models.
[18] Maps corresponding to other parameters are available from the authors upon request
[19] Other variables such as the number of labor hours or the total acreage of the farm have been tested but none of them were significantly correlated to any of the individual parameters.

enough machinery to easily adjust their acreages. Different explanations can lie at the root of the positive and significant correlations between root crop acreage and the $\beta_{k,i}$ and α_i parameters: root crops are good preceding crops for wheat and other cereals which explains the positive correlation with their potential yields; furthermore, a good soil quality is necessary to grow root crops and this good quality also benefits to other crops like wheat and other cereals but also oilseeds, hence the positive correlation with all the $\beta_{k,i}$ parameters; finally, root crops can be used as an alternative to oilseeds as preceding crops for wheat and other cereals and thus relax some constraints on acreage adjustments which translates into a positive correlation with the α_i parameters. The positive and significant correlations between farmers' age and potential yields might be due the role played by experience in farmers' skills and abilities.

All the aforementioned exogenous variables could thus help controlling for part of farm heterogeneity in our production choice model. However, none of the correlations presented in Table is high enough to conclude that using these control variables would be sufficient to capture all the sources of heterogeneity.

4.4.Simulation Results

This last subsection is devoted to the presentation of some simulation results: we simulate the impacts of changes in crop prices corresponding to those that have been observed in France since 2007, namely a 20% in wheat and other cereal prices and a 50% increase in oilseeds prices.

As mentioned in section 4.2, the $\boldsymbol{\gamma}$ and $\boldsymbol{\delta}$ parameters representing the effects of price and time on yields are not very precisely estimated. Therefore, we focus here on the impacts of price changes on acreages and assume that these shocks do not impact yields, which are thus held constant in the simulations.

Table 5 reports the distribution characteristics of the elasticities of acreages to changes in crops prices in our sample. These elasticities are key parameters determining farmers' responses to price shocks. We can first notice that all these calibrated elasticities have the expected signs: own price elasticities are positive and cross price elasticities are negative. They also lie in a reasonable range and reflect the higher flexibility of acreage adjustments within the cereal nest: wheat (respectively other cereals) acreage responds more to a change in other cereals (respectively wheat) price than to a change in oilseed price. Furthermore, the reported quantiles values reflect a great dispersion of elasticities within our sample. One can thus expect each farmer to react differently to the price changes we simulate here, which is not surprising given the variances of the model random parameters.

The first column of Table 6 reports the effects on acreages of the changes in crop prices simulated using our "statistically calibrated" individual parameters model. The relative increase of oilseeds price compared to wheat and other cereals prices lead farmers to reallocate part of their land to this now more profitable crop: among the 10168 ha devoted to crops in our sample, 159 ha of wheat (representing 4% of the initial wheat acreage) and 183 ha of other cereals (representing 6% of the initial other cereals acreage) are reallocated to oilseeds which acreage thus increases by 342 ha (representing 12% of the initial oilseeds acreage). This represents average variations of 2 ha, 2 ha and 4 ha for respectively wheat, other cereals and oilseeds acreage. However, these variations greatly vary from one farm to another: the increases in oilseeds acreage notably vary between 1 ha and 13 ha in absolute term, and between 3% and 38% of the initial oilseeds acreage, depending on the farm. These contrasting results come to illustrate the heterogeneity in farmers' response to economic incentives.

In order to further assess the potential impacts of the approach proposed here to account for heterogeneity on the overall simulated effects of price changes, two alternative versions of the NMNL model have been estimated and used to simulate the same shock. In the first model, all parameters are fixed. This model is estimated using a Maximum Likelihood approach. In the second model, the α and ρ parameters, representing the flexibility of acreages adjustment, are fixed, the $\boldsymbol{\beta}_i$, and $\boldsymbol{\zeta}_i$ are random. This last model can thus be considered as fixed individual effect model. It is estimated using the SAEM algorithm. The estimation results of these two models are presented in Table 7. The main elements that come out of these results are that (*i*) we encounter the same problem to indentify the price effects as with the random parameter model: here the $\boldsymbol{\gamma}$ and $\boldsymbol{\delta}$ parameters are even not significantly estimated; (*ii*) in the fixed effect model, the estimated values of α and ρ are closed to estimated their expectation values in the random parameter model ($\hat{\alpha} = 0.100$ and $\hat{\rho} = 0.122$), which is not the case with the fixed parameter model where $\hat{\alpha}$ and $\hat{\rho}$ respectively equal to 0.017 and 0.045; (*iii*) the log likelihood of the fixed effect and fixed parameter model respectively equal to -1005.4 and -940.27, compared to -816.35 for the random parameter model: the log likelihood ratio test thus clearly indicates that the random parameter model significantly better fits the data than the fixed parameter model ($\chi^2 = 378$, DF=28) and the fixed effect model ($\chi^2 = 248$, DF=13) at the 1% level.

The impacts of price changes on acreages simulated with these two models are reported in the second and third column of Table 6. The overall impacts on acreages are clearly underestimated with the fixed parameter model: the changes in wheat, other cereals and oilseeds acreages are respectively equal to -44 ha, -3 ha and +48 ha, which represent 72% to 98% lower effects than the ones simulated with random parameter model. This can certainly essentially be attributed to the lower estimated values of α and ρ. However, despite $\hat{\alpha}$ and $\hat{\rho}$ values close to their “expectation equivalent” in the fixed effect model, overall simulated impacts also tend to be underestimated in this model, even if to a lesser extent (2% to 40% lower effects). One possible explanation is that farmers owning more land are also the ones that have the more flexibility in acreage adjustments, hence the largest simulated global acreage variations when the heterogeneity in α and ρ is taken into account. These results are clearly illustrated on Figure 4 which reports the individual simulated effects on oilseeds acreage using the three models and taking the random parameter model as reference: the higher the impacts on oilseeds acreages are, the more they are underestimated by the two alternative models. There is thus a risk, by partially or totally ignoring the heterogeneous determinants of farmers’ behaviors in micro econometric models, to generate biased simulation results.

Conclusion

Many unobserved heterogeneous factors can impact farmers' production decisions. The approach we propose in this paper allows accounting for this heterogeneity in the econometric estimations of agricultural production models in a fairly flexible way. We rely on a random parameter modeling framework: the distribution of the model parameters across the farmer population is estimated, which allows the parameters to be farmer specific in order to account for unobserved heterogeneity effects.

Using specific estimators and optimization procedures designed by statisticians, we are able to estimate a random version of the multicrop econometric model proposed by Carpentier and Letort (2013). This empirical application is based on a sample of French crop producers observed from 2004 to 2007. We find that the key parameters of the model exhibit significant variability across farmers. Furthermore, our random model proves to better fit the data than its counterpart fixed or "quasi-fixed" versions. We thus find evidence that heterogeneity significantly matters for the modeling of micro-economic agricultural production choices.

We also show how random parameter models can be used to "statistically calibrate" a simulation model based on a sample of heterogeneous farms and use this "calibrated" model to simulate the impact of crop price changes on acreages. This allows us to further illustrate the potential role of heterogeneity in micro econometric production choices models, and to show that ignoring it can lead to misleading simulation results.

Tables and Figures

Table 1: Selected parameter estimates, yield equation (standard deviations in parentheses)

	γ_k	$E[\beta_{k,i}]$	$Cov[\beta_{k,i},\beta_{l,i}]$			$Var[v_{k,it}]$
			Wheat ($l=1$)	**Cereals** ($l=2$)	**Oilseeds** ($l=0$)	
Wheat ($k=1$)	0.710	7.952	0.992	0.807	0.555	0.595
	(0.120)	(0.051)	(0.077)	(0.073)	(0.046)	(0.044)
Cereals ($k=2$)	0.140	7.167	0.807	1.215	0.509	1.077
	(0.101)	(0.056)	(0.073)	(0.098)	(0.048)	(0.079)
Oilseeds ($k=0$)	0.174	5.265	0.555	0.509	0.428	0.852
	(0.110)	(0.034)	(0.046)	(0.048)	(0.034)	(0.062)

Note: standard errors are in parentheses

Table 2: Selected parameter estimates, acreage share equation (standard deviations in parentheses)

	Expectation	**Covariances with**				
		$\ln\alpha_i$	$\ln\rho_i$	$\ln\beta_{1,i}$ **Wheat**	$\ln\beta_{2,i}$ **Cereals**	$\ln\beta_{0,i}$ **Oilseeds**
$\ln\alpha_i$	-2.357	0.196	0.112	0.008	0.019	0.010
	(0.023)	(0.015)	(0.014)	(0.003)	(0.004)	(0.003)
$\ln\rho_i$	-2.186	0.112	0.279	-0.012	0.007	-0.005
	(0.027)	(0.014)	(0.022)	(0.004)	(0.005)	(0.004)

Note: standard errors are in parentheses

Table 3: Fitting criteria of the model

	Yields y_{kit}			**Acreage shares** S_{kit}		
	"pseudo R^2 "	Observed average	Predicted average	"pseudo R^2 "	Observed average	Predicted average
Wheat ($k=1$)	73.80%	7.93	7.91	79.89%	0.45	0.46
Cereals ($k=2$)	65.38%	7.78	7.74	84.79%	0.30	0.30
Oilseeds ($k=0$)	51.66%	5.73	5.72	56.56%	0.24	0.25

Table 4: Correlations between random parameter values and farmers' characteristics

	$\beta_{k,i}$ Wheat	$\beta_{k,i}$ Other cereals	$\beta_{k,i}$ Oilseeds	α_i	ρ_i
Farm capital	0.172	0.126	0.262	0.165	0.003
	(0.079)	(0.200)	(0.007)	(0.092)	(0.973)
Root crop acreage	0.310	0.195	0.296	0.407	-0.027
	(0.001)	(0.059)	(0.002)	(<0.001)	(0.781)
Farmer's age	0.308	0.206	0.282	0.157	-0.242
	(0.001)	(0.035)	(0.004)	(0.111)	(0.013)

Note: P-values are in parentheses

Table 5: Characteristics of the distribution of acreage shares price elasticities

	Average	Q5	Q25	Q50	Q75	Q95
Wheat Acreage						
Wheat Price	0.43	0.24	0.32	0.39	0.49	0.77
Other cereals Price	-0.25	-0.61	-0.29	-0.18	-0.15	-0.11
Oilseeds Price	-0.14	-0.24	-0.16	-0.13	-0.11	-0.08
Other cereals acreage						
Wheat Price	-0.48	-1.19	-0.67	-0.36	-0.23	-0.14
Other cereals Price	0.61	0.22	0.33	0.49	0.79	1.33
Oilseeds Price	-0.14	-0.24	-0.16	-0.13	-0.11	-0.08
Oilseeds acreage						
Wheat Price	-0.37	-0.84	-0.45	-0.31	-0.20	-0.13
Other cereals Price	-0.23	-0.67	-0.29	-0.17	-0.10	-0.05
Oilseeds Price	0.50	0.17	0.31	0.43	0.65	0.95

Table 6: Simulated impacts on acreages of the price shock

	Random parameter model	Fixed parameter model	Fixed Individual effects model
Wheat Acreage			
Total change in ha	-159 (-3.9%)	-44 (-0.9%)	-96 (-2.2%)
Average change in ha	-2 (-4.5%)	-1 (-0.9%)	-1 (-2.3%)
Max change in ha	<0.5 (<0.1%)	1 (+0.6%)	1 (+1%)
Min change in ha	-7 (-17.0%)	-2 (-2.4%)	-5 (-9.3%)
Other cereals Acreage			
Total change	-183 (-5.6%)	-3 (-0.1%)	-178 (-4.7%)
Average change	-2 (-6.3%)	<0.5 (-0.1%)	-2 (-5.1%)
Max change	+2 (+11.7%)	1 (+1.8%)	<0.5 (<0.1%)
Min change	-8.8 (-20.0%)	-1 (-2.7%)	-8.8 (-20.0%)
Oilseeds acreage			
Total change	+342 (+12.1%)	+48 (+2.4%)	+274 (+12.8%)
Average change	+4 (+13.9%)	+1 (+2.4%)	+3 (+12.8%)
Max change	+13 (+38.2%)	+1 (+4.0%)	+9 (+21.8%)
Min change	1 (+3.1%)	<0.5 (+1.1%)	+1 (+4.5%)

Note: Numbers in parentheses correspond to %age changes compared to initial acreages

Table 7: Results of the three models

Parameters	Model1 (All parameters are fixed)	Model2 (α and ρ are fixed)	Model3 (Random parameters Model)
$\ln\beta_1$	1.972 (0.008)	2.085 (0.004)	2.066 (0.006)
$\ln\beta_2$	2.123 (0.009)	1.956 (0.005)	1.958 (0.007)
$\ln\beta_0$	1.645 (0.011)	1.579 (0.004)	1.653 (0.006)
$\ln\alpha$	-4.502 (0.021)	-2.303 (0.033)	-2.357 (0.023)
$\ln\rho$	-3.351 (0.033)	-2.104 (0.019)	-2.186 (0.027)
ζ_1	-94.887 (3.072)	-4.431 (0.098)	-4.015 (0.141)
ζ_2	-77.982 (3.380)	-1.752 (0.152)	-1.722 (0.130)
γ_1	-0.562 (0.091)	0.942 (0.138)	0.710 (0.120)
γ_2	1.453 (0.095)	0.133 (0.114)	0.140 (0.101)
γ_0	-0.045 (0.109)	-0.456 (0.130)	0.174 (0.111)
$\delta_{1,2004}$	1.131 (0.137)	1.052 (0.155)	1.078 (0.129)
$\delta_{1,2005}$	0.247 (0.106)	0.292 (0.161)	0.296 (0.151)
$\delta_{1,2006}$	-0.050 (0.133)	0.190 (0.130)	0.154 (0.121)
$\delta_{2,2004}$	0.863 (0.143)	1.253 (0.160)	1.248 (0.135)
$\delta_{2,2005}$	0.644 (0.140)	0.764 (0.253)	0.767 (0.162)
$\delta_{2,2006}$	0.458 (0.156)	0.422 (0.125)	0.429 (0.132)
$\delta_{0,2004}$	1.061 (0.115)	1.163 (0.146)	1.078 (0.131)
$\delta_{0,2005}$	0.816 (0.103)	0.885 (0.130)	0.844 (0.180)
$\delta_{0,2006}$	0.031 (0.09)	0.043 (0.184)	0.047 (0.147)
$\Lambda_{1,1}$	1.496 (0.116)	0.587 (0.032)	0.595 (0.044)
$\Lambda_{1,2}$	0.945 (0.102)	0.213 (0.041)	0.217 (0.043)
$\Lambda_{1,3}$	0.704 (0.077)	0.153 (0.042)	0.162 (0.038)
$\Lambda_{2,2}$	2.188 (0.168)	1.074 (0.061)	1.077 (0.079)
$\Lambda_{2,3}$	0.647 (0.091)	0.143 (0.058)	0.159 (0.050)
$\Lambda_{3,3}$	1.277 (0.114)	0.831 (0.060)	0.852 (0.062)
$\Omega_{1,1}$	---	0.0156 (0.0007)	0.016 (0.0012)
$\Omega_{1,2}$	---	0.0143 (0.0009)	0.014 (0.0012)
$\Omega_{1,3}$	---	0.0138 (0.0007)	0.013 (0.0010)
$\Omega_{1,4}$	---	---	0.008 (0.0029)
$\Omega_{1,5}$	---	---	-0.008 (0.0037)
$\Omega_{1,6}$	---	-0.1559 (0.0123)	-0.084 (0.0181)
$\Omega_{1,7}$	---	-0.1060 (0.0195)	0.027 (0.0157)

$\Omega_{2,2}$	---	0.0239 (0.0013)	0.023 (0.0017)
$\Omega_{2,3}$	---	0.0148 (0.0009)	0.013 (0.0012)
$\Omega_{2,4}$	---	---	0.019 (0.0038)
$\Omega_{2,5}$	---	---	0.007 (0.0045)
$\Omega_{2,6}$	---	-0.1738 (0.015)	-0.062 (0.0221)
$\Omega_{2,7}$	---	-0.1488 (0.022)	0.010 (0.0191)
$\Omega_{3,3}$	---	0.0179 (0.00083)	0.015 (0.0012)
$\Omega_{3,4}$	---	---	0.010 (0.0030)
$\Omega_{3,5}$	---	---	-0.005 (0.0037)
$\Omega_{3,6}$	---	-0.1668 (0.0134)	-0.088 (0.0184)
$\Omega_{3,7}$	---	-0.0543 (0.0202)	0.029 (0.0159)
$\Omega_{4,4}$	---	---	0.196 (0.0147)
$\Omega_{4,5}$	---	---	0.112 (0.0142)
$\Omega_{4,6}$	---	---	0.062 (0.0605)
$\Omega_{4,7}$	---	---	-0.126 (0.0543)
$\Omega_{5,5}$	---	---	0.279 (0.0217)
$\Omega_{5,6}$	---	---	-0.409 (0.0802)
$\Omega_{5,7}$	---	---	0.421 (0.0721)
$\Omega_{6,6}$	---	7.767 (0.417)	7.408 (0.5327)
$\Omega_{6,7}$	---	-6.0536 (0.573)	-5.709 (0.4352)
$\Omega_{7,7}$	---	19.349 (0.997)	6.239 (0.4150)
Log Likelihood	-1005.356	-940.270	-816.352
Likelihood ratio test: H0: Model1	---	130.172 DF=15	378.008 DF=28
Likelihood ratio test: H0: Model2	---	---	247.836 DF=13

Note: standard errors are in parentheses

Figure 1: ***Ex post*** **and** ***ex ante*** **distribution of random parameters**

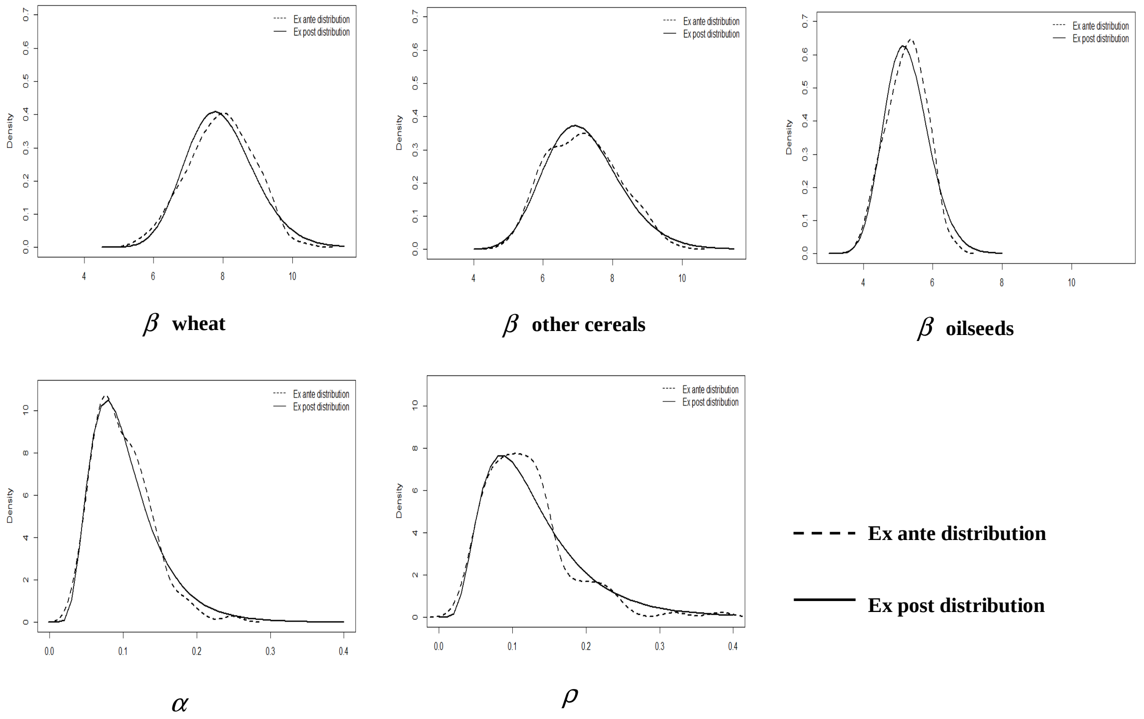


Figure 2: Calibrated values and confidence intervals of individual parameters

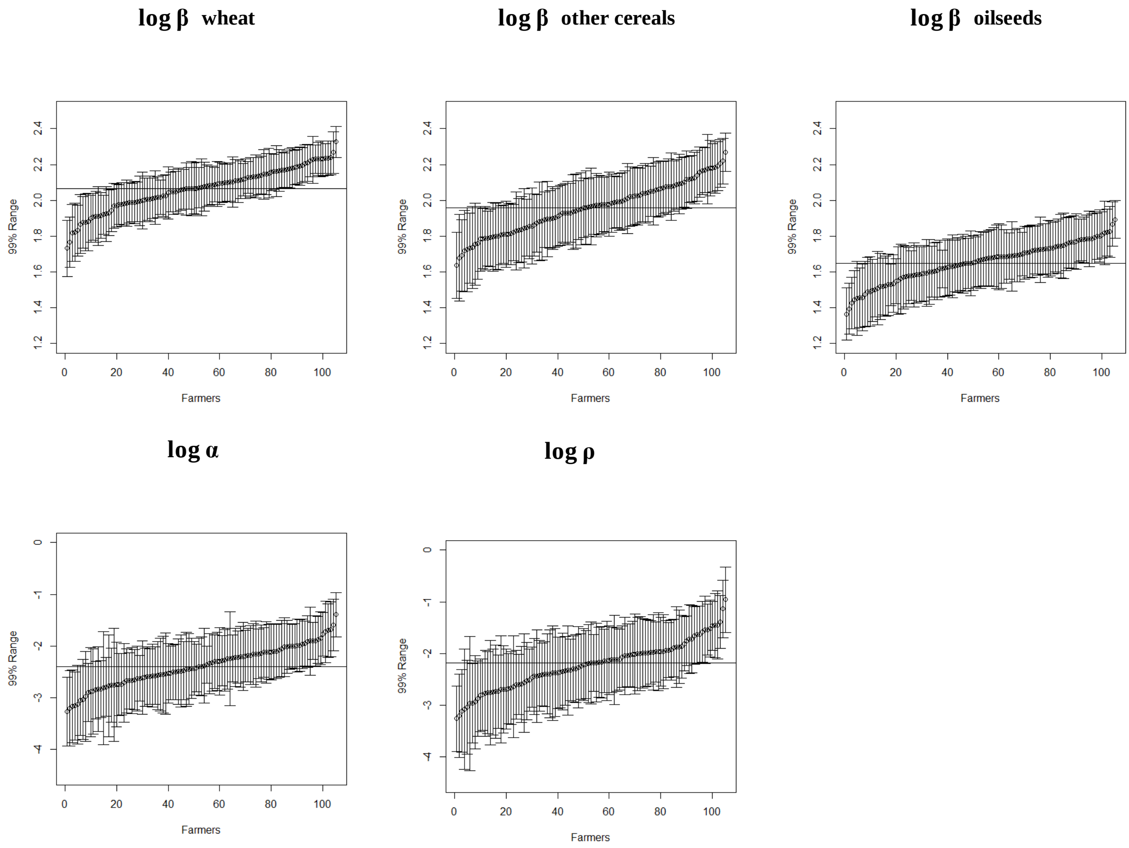
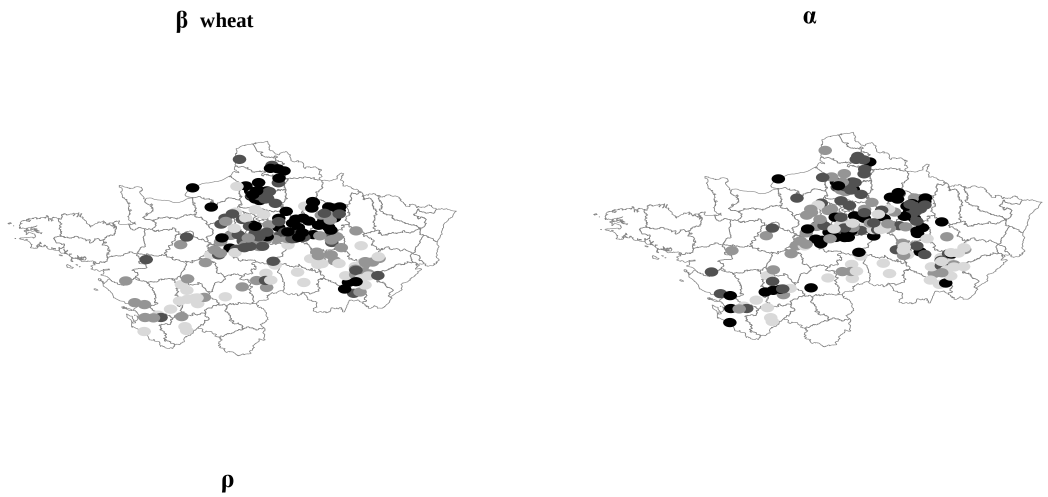


Figure 3: Distribution of selected random parameters across the population sample



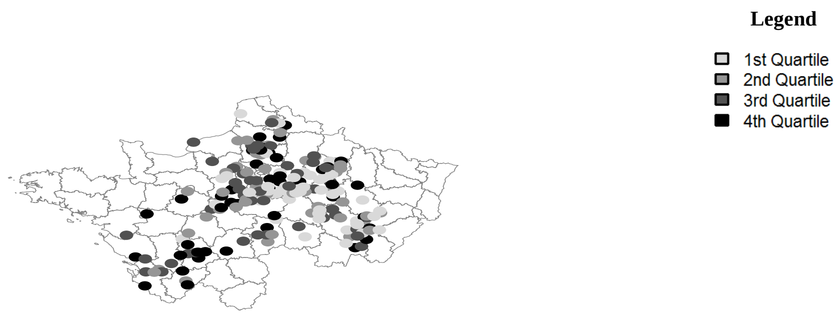
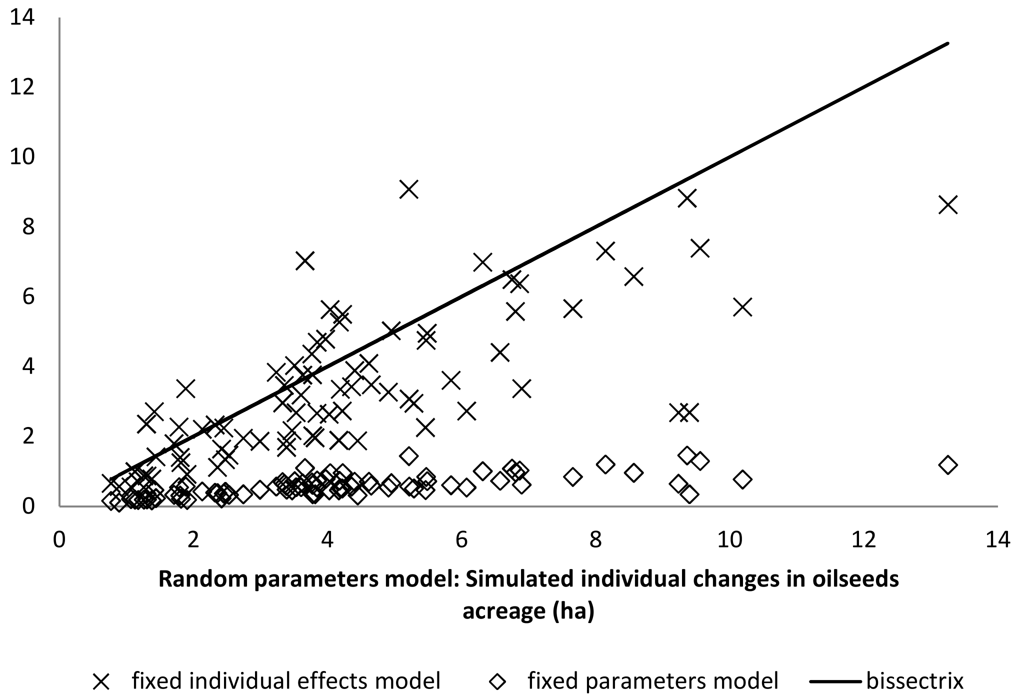


Figure 4: Comparison of the impacts on oilseeds acreage simulated with the different models



References

Ackerberg. D., Benkard. L., Berry S., Pakes A., 2007. Econometric Tools for Analyzing Market Outcomes. In J.J. Heckman and E.E. Leamer. eds., *Handbook of Econometrics*. vol. 6A. Amsterdam. Elsevier.

Berry. S. T., 1994. Estimating Discrete-Choice Models of Product Differentiation. *RAND Journal of Economics.* 25, 242-262.

Caffo, B. S., Jank W., Jones G. L., 2005. "Ascent-based Monte Carlo expectation–maximization." *Journal of the Royal Statistical Society: Series B (Statistical Methodology).* 67.2, 235-251.

Caponi V., 2011. Intergenerational transmission of abilities and self-selection of mexican immigrants. *International Economic Review.* 52(2), 523-547.

Carpentier A., Letort E., 2013. Multicrop production models with Multinomial Logit acreage shares. *Environmental and Resource Economics.* (2013), 1-23.

Carpentier A., Letort E., 2012. "Accounting for heterogeneity in multicrop micro-econometric models: Implications for variable input demand modelling." *American Journal of Agricultural Economics.* 94(1), 209–224.

Chambers R., Just R., 1989. Estimating Multiouput Technologies. *American Journal of Agricultural Economics.* 71, 980-95.

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

Delyon B., Lavielle M., Moulines E., 1999. Convergence of a stochastic approximation version of the EM algorithm. *Annals of Statistics*. 94-128.

Dempster A. P., Laird N. M., Rubin D. B., 1977. Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society Series B (Methodological)*, 1-38.

Eaton J., Kortum S., Kramarz F., 2011. An Anatomy of International Trade: Evidence From French Firms. *Econometrica.* 79(5), 1453-1498.

Heckeleï T., Britz W., Zhang Y., 2012. Positive mathematical programming approaches. Recent developments in literature and applied modelling. *Bio-based and Applied Economics.* 1(1),109-124.

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

Heckman J. J., 2001. Micro data. heterogeneity. and the evaluation of public policy: Nobel lecture. *Journal of political Economy*. 109(4). 673-748.

Heckman J. J., Sedlacek G., 1985. Heterogeneity. aggregation. and market wage functions: an empirical model of self-selection in the labor market. *The Journal of Political Economy*. 1077-1125.

Howitt E., 1995. Positive Mathematical Programming. *American Journal of Agricultural Economics*. 77,329-342.

Jank W., Booth J., 2003. Efficiency of Monte Carlo EM and Simulated Maximum Likelihood in Tow-Stage Hierarchical Models. *Journal of Computational and Graphical Statistics.* 12(1), 214-229.

Keane M. P., 2009. Simulated Maximum likelihood estimation based on FOCs. *International Economic Review*. 50(2), 627-675.

McFadden D., Train. K., 2000. Mixed MNL models for discrete response. *Journal of applied Econometrics*. 15(5), 447-470.

McLachlan G., Krishnan T., 2008. *The EM algorithm and extensions*. 2nd edition. Wiley Edition.

Meng X. L., Rubin. D. B., 1993. Maximum likelihood estimation via the ECM algorithm: A general framework. *Biometrika.* 80(2), 267-278.

Moro D., Sckokai P. 2013. The impact of decoupled payments on farm choices: Conceptual and methodological challenges. *Food Policy*. 41:28-38.

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

Train K., 2009. *Discrete Choice Methods with Simulation* (2nd ed). Cambridge: University Press.

Train K., 2008. EM algorithms for nonparametric estimation of mixing distributions. *Journal of Choice Modelling.* 1(1), 40-69.

Train K., 2007. A *recursive estimator for random coefficient models.* University of California. Berkeley.

Wei G. C., 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(411), 699-704.