Modelling Acreage Decisions within the Multinomial Logit Framework:

Profit Functions and Discrete Choice Models

A. Carpentier* and E. Letort**
***INRA, Rennes and ENSAI, Rennes**
**** INRA, Rennes**

Selected Paper prepared for presentation at the American Agricultural Economics Association Annual meeting, Orlando, FL, July 27-29, 2008.

1. Introduction

The Multinomial Logit (MNL) framework developed by McFadden (see, *e.g.*, Train, 2003) was used by several authors (see, *e.g.*, Lichtenberg, 1989; Wu and Segerson, 1995) to model acreage decision. It was also used to represent input cost shares in a static and in dynamic framework (see, *e.g.*, Considine and Mount, 1984 and Jones, 1995). The models built within this framework are mainly used for three reasons: they define shares strictly lying between 0 and 1, they are flexible and their parameters are easy to estimate thanks to transformation of the share equations into linear Logit models.

Nevertheless, in the context of acreage allocation modelling, the MNL framework was used more as a statistical modelling framework than as a theoretically grounded model. Furthermore, the MNL framework has only been used to define standard Logit share models in a static framework.

The first purpose of this paper is to propose theoretical justifications for using Logit acreage share models. Two approaches are presented: the Logit shares can be derived from a well defined profit function or derived as the result of a set of discrete choices. It is next shown that both theoretical frameworks allow to define generalizations of the standard Logit shares. These generalizations build on developments of the MNL Logit framework for modelling discrete choices and seek to define models that are flexible and empirically tractable.

The main approach used by agricultural economists to define land allocation models is to derive them from profit or restricted profit functions (see, *e.g.*, Chambers and Just, 1989). Building on the work of Anderson *et al.* (1992) it is shown that Logit acreage shares can be derived from a specific restricted profit function. This profit function has a structure similar to the profit function used to derive land allocation functions when the multicrop production is nonjoint in outputs. It is defined as the sum of crop specific profit functions that are weighted by the acreage shares (see, *e.g.*, Chambers and Just, 1989). Thus, it is defined as the weighted sum of the crop gross margins minus an implicit cost function of the land allocation. This cost function is defined by the the (cross-)entropy of the acreage shares. This implicit cost is maximum (and equal to 0) if the acreage shares are equal to some "reference" allocation of the total area. It can be interpreted as the implicit cost of managing the chosen allocation (work peak loads, …). This profit function thus explicitly defines a trade-off between the crop gross return of the different crops and the cost of managing the chosen allocation of the land. The paper also presents the restricted profit function from which Nested Logit acreage share functions can be derived.

Logit acreage shares can also be defined as the result of plot by plot discrete decisions, along the lines of Caswell and Zilberman (1985). This approach is not structural but is rather flexible and allows to consider some dynamic aspects of the acreage decisions.

The farmer is assumed to have *N* plots of equal size and to decide for each which crop to grow. If the profit function for the different crops exhibits constant returns to land area and if the random term (unobserved by the econometrician but observed by the farmers) are independent across plots and identically distributed with the Gumbel distribution, the expected share of land devoted to a given crop has a Logit form. The main drawback of this approach is to consider that the decisions for the different plots are independent. This model can be generalized in two ways.

The first is to define a partial adjustment model where the Logit shares are the "target" (or long run equilibrium) land shares. This allows to account for the adjustment costs of the land allocation decisions. This results in nonlinear (in parameters) but tractable share equations. This model can be generalized to explicitly consider the effects of crop rotations on profits on the one hand and their effect on farmers' choices. In particular, this approach allows the use of the results derived by the literature initiated by Rust (1987) for the analysis of dynamic discrete choices.

The rest of the paper is divided in four sections. The second and third sections present the two approaches that can lead to use of the MNL framework to model acreage shares. They detail specification and statistical inference issues and stress the main advantages and drawbacks of the presented approaches.

Two applications are presented in the fourth part to illustrate the empirical interest of the proposed models. Both use a rotating panel of French farms (1987-2006) and consider the estimation of yield functions, variable input demand functions and acreage share functions.

The last section concludes and suggests possible extensions.

2. Acreage decisions within the MNL framework: profit functions

The typical short run problem faced by a farmer is to allocate his land to K different crops ($k=1,...,K$) according to the acreage shares s_k for $k=1,...,K$ with $\sum_{k=1}^{K} s_k = 1$. Crop k output is sold at price p_k and uses short run inputs in quantity $\mathbf{x}_k$ that are bought at prices **w**. The derivation of MNL acreage share functions rely on some primary assumptions related to the multicrop technology. These assumptions and their main implications are presented in the following paragraph. They are discussed in more details at the end of this section.

Let assume that the multicrop technology is non joint in outputs in the short run and exhibits constant returns to land on a crop *per* crop basis. These assumptions are discussed later in this section. Along with the assumption that the short run production technology of each crop k is concave in $\mathbf{x}_k$ these assumptions imply that well behaved short run profit functions *per* unit of land exist for each crop. These profit functions, by assumption, do not depend on **s**, the vector of acreage shares. It is here further assumed the short run inputs and the fixed inputs used by the farmer are "separable" in the multicrop technology in the sense that the fixed factors does not affect the short run crop specific production technologies whereas it affects the acreage choices possibilities. More specifically, it is assumed that the restricted (in **s**) short run profit function (*per* unit of land) associated with the considered multicrop technology has the following form:

$$(1) \qquad \Pi(\mathbf{p},\mathbf{w};\mathbf{s}) \equiv \sum_{k=1}^{K} s_k \pi_k(p_k,\mathbf{w}) - C(\mathbf{s})$$

where the $\pi_k(p_k,\mathbf{w})$ term denotes the short run profit function of crop k and $C(\mathbf{s})$ denotes the cost function associated to the choice of acreage shares denoted by **s**. According to these assumptions, the forms of the crop specific restricted short run functions $\pi_k(p_k,\mathbf{w})$ and of the cost function $C(\mathbf{s})$ depend on the available (quasi-)fixed production factors. However, these assumptions imply that the considered farmer short run decisions for each crop, *i.e.* target yield and short run input uses, do not depend on his acreage choices.

It is important to stress the fact that the profit function defined in (1) is only intended to represent the short run objective function of the farmer. Farmers adapt their short run input and their acreage choices in the short run, but their adaptation possibilities are limited in the short run. In this context the cost function $C(\mathbf{s})$ is a reduced form function representing the effects of the constraints and costs faced by the farmer for the choice of his land allocation in the short run. The form of this cost function depends on the available quantities of (quasi-)fixed production factors, these quantities defining the existence of work or machinery peak loads. In the considered static framework this "acreage management cost function" is the main motive of crop diversification of the farmer.

In this simple static context, the acreage shares chosen by the considered farmer have a MNL form if we have:

(2) $$s_k(\mathbf{p},\mathbf{w}) = \frac{\exp\left[\dfrac{\pi_k(p_k,\mathbf{w})-c_k}{a}\right]}{\sum_{m=1}^{K}\exp\left[\dfrac{\pi_m(p_m,\mathbf{w})-c_m}{a}\right]} = \frac{\exp\left[\dfrac{\left(\pi_k(p_k,\mathbf{w})-\pi_K(p_K,\mathbf{w})\right)-(c_k-c_K)}{a}\right]}{1+\sum_{m=1}^{K-1}\exp\left[\dfrac{\left(\pi_m(p_m,\mathbf{w})-\pi_K(p_K,\mathbf{w})\right)-(c_m-c_K)}{a}\right]},$$

where $\mathbf{p}$ is the vector of output prices. The terms a and the c_k are parameters (that can be expressed as functions of the characteristics of the farmer and of its farm) to be defined determining the land allocation. In the following the c_k parameters are assumed to be strictly positive, allowing their interpretation as fixed costs per unit of land devoted to crop *k*.

The models built within this framework have two main advantages: it defines shares strictly lying between 0 and 1 and, assuming that the crop specific profit functions are known or can be estimated, their parameters are easy to estimate thanks to transformation of the share equations into "linear" Logit models:

(3) $$\ln s_k(\mathbf{p},\mathbf{w}) - \ln s_K(\mathbf{p},\mathbf{w}) = a^{-1}\left[\pi_k(p_k,\mathbf{w}) - \pi_m(p_m,\mathbf{w}) - (c_k - c_K)\right], \quad k = 1,\dots,K-1.$$

As will be shown below, the share functions have a third main advantage: they can be derived from a profit function that is easily interpretable as the objective function for short run land allocation decision. This MNL acreage shares can be derived from the following restricted profit function:

(4) $$\Pi\left[\mathbf{s},\boldsymbol{\pi}(\mathbf{p},\mathbf{w})\right] \equiv \sum_{k=1}^{K} s_k\pi_k(p_k,\mathbf{w}) - \left[A + a\sum_{k=1}^{K} s_k(c_k + \ln s_k)\right]$$

where $\boldsymbol{\pi}$ denotes the vector of the $\pi_k(p_k,\mathbf{w})$'s. Building on the work of Anderson *et al.* (1992) it is easily shown that the maximisation of (4) according to $\mathbf{s}$ subject to $\sum_{k=1}^{K} s_k = 1$ (and $\mathbf{s} > \mathbf{0}$) leads to acreage share functions of the form defined in (2). The restricted profit function (4) allows to define the associated cost function:

(5) $$C(\mathbf{s}) = A + a\sum_{k=1}^{K} s_k(c_k + \ln s_k) = A + a\sum_{k=1}^{K} s_k c_k + a\sum_{k=1}^{K} s_k \ln s_k \ .$$

This cost function is defined is composed of three kinds of terms: a (unidentifiable) fixed cost: A, *per* unit of land crop specific fixed costs: c_k for $k = 1,...,K$ (only the differences $(c_k - c_R)$ for $k = 1,...,K$ and $k \neq R$ can be identified, R denoting the reference crop) and the (opposite of the) entropy function of the acreage defined as land shares: $\sum_{k=1}^{K} s_k \ln s_k$. This function is negative and its minimum value is achieved in $s_k = K^{-1}$ for $k = 1,...,K$ (the term A can be chosen to ensure that the cost function is positive). It implies that crop diversification reduces of management cost. This cost function is easier to interpret by using an alternative but equivalent specification. If the c_k are defined as:

(6) $$c_k = -\ln d_k \quad \text{with} \quad \sum_{k=1}^{K} d_k = 1$$

(and the A parameter is modified correspondingly to A_d), the part of $C(\mathbf{s})$ that depends on s can be defined as:

(7) $$a\sum_{k=1}^{K} s_k(\ln s_k + c_k) = a\sum_{k=1}^{K} s_k(\ln s_k - \ln d_k) \ .$$

The term $\sum_{k=1}^{K} s_k(\ln s_k - \ln d_k)$ is the (opposite) of the cross-entropy function of the land allocation shares. Its minimum is achieved where $s_k = d_k$ for $k = 1,...,K$. This leads to the interpretation of **d**, the vector of the d_k', as a reference land allocation in terms of management costs. It defines the land allocation for which the management costs are minimum, *i.e.* the acreage that is the most suitable to the farm, and in particular its (quasi-)fixed inputs. Any acreage **s** different from **d** is more costly to manage than **d**, and the more the acreage **s** is different form **d** according to the distance function defined by the cross-entropy function, the more **s** is costly.

Omitting the A term (as will be the case in what follows), the restricted profit function (4) can now be rewritten as:

$$\Pi\left[\mathbf{s},\boldsymbol{\pi}(\mathbf{p},\mathbf{w})\right]=\sum_{k=1}^{K}s_k\pi_k(p_k,\mathbf{w})-a\sum_{k=1}^{K}s_k(\ln s_k-\ln d_k)\,. \tag{8}$$

This specification provides the interpretation of the a term: this term defines the weight of the of the management costs. The cost associated to the choice of acreage shares different from $\mathbf{d}$ increases as a^{-1} increases. In particular it can be shown that:

$$\begin{aligned}&s_k(\mathbf{p},\mathbf{w};a,\mathbf{d})\xrightarrow[a\to 0]{}1 \text{ if } \pi_k(p_k,\mathbf{w})=Max\{\pi_m(p_m,\mathbf{w});\quad m=1,...,K\},\\&s_k(\mathbf{p},\mathbf{w};a,\mathbf{d})\xrightarrow[a\to 0]{}0 \text{ otherwise}\end{aligned} \tag{9a}$$

and:

$$s_k(\mathbf{p},\mathbf{w};a,\mathbf{d})\xrightarrow[a\to +\infty]{}d_k\,. \tag{9b}$$

The farmer only grows the most profitable crop tends to 0 and he chooses the minimum cost acreage $\mathbf{d}$ if a tends to infinity.

In fact, this framework is suitable for modelling moderate acreage modifications. It can be interpreted as local approximation of the "true" objective function for acreages shares belonging to a neighbourhood of $\mathbf{d}$. It is empirically tractable but the heterogeneity of the farmers and of their farms requires the parameters a and $\mathbf{d}$ to be defined as functions accounting for this heterogeneity.

Another advantage of this framework is that is allows to define the associated profit function. It has the well-known log-sum form:

$$\Pi^*\left[\boldsymbol{\pi}(\mathbf{p},\mathbf{w});a,\mathbf{c}\right]\equiv a\ln\left[\sum_{k=1}^{K}\exp\left[\frac{\pi_k(p_k,\mathbf{w})-c_k}{a}\right]\right] \tag{10a}$$

or:

(10b) $\Pi^*\left[\boldsymbol{\pi}(\mathbf{p},\mathbf{w});a,\mathbf{d}\right] \equiv a\ln\left[\sum_{k=1}^{K} d_k \exp\left[\pi_k(p_k,\mathbf{w})\right]\right].$

The main drawback of these cost and profit functions is that they consider all crops as equivalently, especially in terms of management costs. But crops can be similar or, at the opposite, very different according to their needs at different points of the growing season. Some crops may also have similar properties in terms of crop rotations. As the standard MNL model used for discrete choices was generalised to account for similarities of the choice alternatives (see, *e.g.*, Train, 2003) this framework can be generalised to account for similarities in the management of different crops. If we assume that the K crops can be allocated to Q mutually exclusive nests, it is possible to define corresponding Nested MNL acreage share functions and their corresponding profit and restricted profit functions. The set of crops belonging to nest q, $q=1,...,Q$, is denoted by $B(q)$, the share of land allocated to the crops of nest q is denoted by $\bar{s}_q = \sum_{k\in B(q)} s_k$ and the share of crop k within its nest, denoted by $q(k)$, is denoted by $s_{kq(k)} = s_k / \bar{s}_{q(k)}$. Building on the work of Verboven (1996), it can be shown that the maximisation in **s** of the restricted profit function:

(11a) $\Pi(\mathbf{s},\boldsymbol{\pi};a,\boldsymbol{\alpha},\mathbf{c}) \equiv \sum_{q=1}^{Q}\sum_{k\in B(q)} s_k(\pi_k - c_k) - a\left[\sum_{q=1}^{Q}\bar{s}_q \ln \bar{s}_q + \sum_{q=1}^{Q}\frac{\alpha_q}{a}\bar{s}_q \sum_{k\in B(q)} s_{kq}\ln s_{kq}\right]$

subject to the total land allocation constraint (and the share positivity constraints) leads to the following Nested MNL acreage share functions:

$$\text{(11b)}\quad s_k(\boldsymbol{\pi};a,\boldsymbol{\alpha},\mathbf{c}) \equiv \frac{\exp\left[\alpha_{q(k)}^{-1}(\pi_m - c_m)\right]}{\sum_{m\in B(q(k))} \exp\left[\alpha_{q(k)}^{-1}(\pi_m - c_m)\right]} \frac{\left[\sum_{m\in B(q(k))} \exp\left[\alpha_{q(k)}^{-1}(\pi_m - c_m)\right]\right]^{\frac{\alpha_{q(k)}}{a}}}{\sum_{q=1}^{Q}\left[\sum_{m\in B(q)} \exp\left[\alpha_{q(k)}^{-1}(\pi_m - c_m)\right]\right]^{\frac{\alpha_q}{a}}}.$$

In the restricted profit function a is the weight parameter of the management cost function for the different nests while α_q is the weight parameter of the management cost function for the crops of nest q. The corresponding management cost function is defined as a sum of the (cross) entropy function of the shares devoted the nests of crops ($\overline{s}_q$) and the weighted (by $\overline{s}_q$) sum of the (cross) of the shares devoted to crops of the Q nests. The form of share functions (11b) is rather specific, it allows to define the share function of crop k within the land allocated to its nest:

$$\text{(11c)}\quad s_{kq(k)}(\boldsymbol{\pi}_{q(k)};\alpha_{q(k)},\mathbf{c}_{q(k)}) \equiv \frac{\exp\left[\alpha_{q(k)}^{-1}(\pi_m - c_m)\right]}{\sum_{m\in B(q(k))} \exp\left[\alpha_{q(k)}^{-1}(\pi_m - c_m)\right]}.$$

This function only depends on the parameters associated to the nest of k, implying some form of weak separability of the shares defined in a given with respect to the shares allocated within the other nests. It is also possible to define the share function of the land allocated to a nest q:

$$\text{(11d)}\quad s_q(\boldsymbol{\pi};a,\boldsymbol{\alpha},\mathbf{c}) \equiv \frac{\left[\sum_{m\in B(q(k))} \exp\left[\alpha_{q(k)}^{-1}(\pi_m - c_m)\right]\right]^{\frac{\alpha_{q(k)}}{a}}}{\sum_{q=1}^{Q}\left[\sum_{m\in B(q)} \exp\left[\alpha_{q(k)}^{-1}(\pi_m - c_m)\right]\right]^{\frac{\alpha_q}{a}}} = \frac{\exp\left[a^{-1}\Pi_q(\boldsymbol{\pi}_q;\alpha_q,\mathbf{c}_q)\right]}{\sum_{\ell=1}^{Q}\exp\left[a^{-1}\Pi_q(\boldsymbol{\pi}_\ell;\alpha_\ell,\mathbf{c}_\ell)\right]}.$$

This share function depends on all parameters of the model. It can also be defined in a standard MNL form by using the $\Pi_q(\boldsymbol{\pi}_q;\alpha_q,\mathbf{c}_q)$ functions. $\Pi_q(\boldsymbol{\pi}_q;\alpha_q,\mathbf{c}_q)$ is defined as the profit functions (the inclusive value) associated to nest q:

(11e) $$\Pi_q(\boldsymbol{\pi}_q;\alpha_q,\mathbf{c}_q)\equiv\alpha_q\ln\left[\sum_{m\in B(q)}\exp\left[\alpha_q^{-1}(\pi_m-c_m)\right]\right].$$

It provides the profit associated to an optimal allocation of the land quantity devoted to the crops belonging to nest q. The profit function associated to the restricted profit function (11a) can also be defined in a standard MNL form by using the $\Pi_q(\boldsymbol{\pi}_q;\alpha_q,\mathbf{c}_q)$ functions:

(11f) $$\Pi^*(\boldsymbol{\pi};a,\boldsymbol{\alpha},\mathbf{c})\equiv a\ln\left[\sum_{\ell=1}^{Q}\exp\left[a^{-1}\Pi_q(\boldsymbol{\pi}_\ell;\alpha_\ell,\mathbf{c}_\ell)\right]\right].$$

The Nested MNL framework can be used to derive interesting results of comparative statics. It less tractable than the standard MNL framework since there no simple counterpart to the log-transformation defined in equation (3). In the particular cases, *e.g.* in cases where the acreage distinguishes a single specific crop with respect to the others (an "outside" crop), the technique developed by Berry (1994) may be used to define empirically tractable estimating equation. The other cases require integrations of the nests profit functions $\Pi_q(\boldsymbol{\pi}_q;\alpha_q,\mathbf{c}_q)$ by simulation techniques (see, *e.g.*, Train, 2003) or approximations.

An empirical illustration of the use of the MNL framework using micro-economic panel data on the French grain crop sector is presented in the fourth section. It uses the standard MNL framework. A Nested MNL version of the model (using the oilseeds and protein crop aggregate as the "outside" crop and Berry's technique) provided comparable results (which are available upon request from the authors).

When compared to the author modelling frameworks used for land allocation choices, the MNL framework developed in this section relies on two major assumptions: non jointness in outputs and constant return to land on a crop *per* crop basis.

Albeit discussed in some contexts, non-jointness in outputs is a common assumption, at least in static frameworks as the one used here (see, *e.g.*, Chambers and Just, 1989).

The constant return to land on a crop *per* crop basis is used some models derived within linear or positive mathematical programming (PMP) frameworks (see the discussion in Howitt, 1995), in models where land allocation choices are derived as a portfolio choices (e.g., Chavas and Holt, 1990 or Coyle, 1992) or in some models built for the allocation of variable inputs (*e.g.*, Just, Zilberman, Hochman and Bar-Shira, 1990). However, agricultural economists modelling static multicrop choices either in a primal (Just, Zilberman and Hochman, 1983) or in a dual approach (Chambers and Just, 1989) usually assume that the gross margin provided by a given crop is a decreasing function of the land quantity devoted to this crop. Howitt (1995) use a similar assumption in a model built to present the PMP methodology. Howitt argues that variability in soil quality implies important heterogeneity in crop yield in most agricultural areas and that crop rotations may induce consequent variability in the suitability of the plots for specific crops. These effects of crop rotations may be approximated by the use of functional forms of gross margins such as the now widely used normalized quadratic profit function suggested in this context by Negri and Moore (1992).

If soil quality may be highly variable at an aggregate level, *e.g.* on the scale of the 21 French (continental) *régions*, this variability is much more limited at the micro-econometric level, *e.g.* on the usual scale of French farms (100 ha on average in our sample). In particular, farms specialized in grain production are mainly located in large fertile plains (large Paris Basin for wheat, Aquitaine for corn,…) characterized by homogenous pedo-climatic conditions. Crop rotations have two sorts of effects. First, they induce constraints on land availability generated by biological cycle. Defining $C(\mathbf{s})$ as a function of the preceding acreage allows to account for such constraints. Second, crop rotations may affect the form of the crop specific yield functions. According to this effect, the assumption that crop specific yields or profit functions are decreasing in land relies on the assumption that profit maximising farmers optimally use the crop rotations according to their properties. In the framework defined in this paper these effects must be account for in the crop specific yields or profit functions themselves. It must be emphasized models of farmer's choices considering the impact of past acreages on current choices must be account for dynamic optimisation by the farmer to be fully consistent.

The specification of dynamic models of land allocation is a difficult task that as only be carried out in specific contexts (see, *e.g.*, Eckstein, 1984 ; Ozarem and Miranowski, 1994 ; Thomas, 2003 or Hennessy, 2007). The applications presented in the paper propose a simple albeit crude approach to this problem.

The next section of the paper presents an alternative modelling framework based on MNL acreage share functions. Contrary to the framework developed in this section, it is not based on a (semi-)structural model of the multicrop production technology. However, the empirical illustration presented in the fourth section yields interesting and promising results. Its simplicity provides some flexibility for further developments.

3. Acreage decisions within the MNL framework: discrete choices

Logit acreage shares can also be defined as the result of plot by plot discrete decisions, along the lines of Caswell and Zilberman (1985). This approach is not structural but appears to be rather flexible. It is based on two main points: the aggregation of choices on a plot *per* plot basis and the logic of partial adjustment of acreage choices. It yields simple models that are empirically tractable and, perhaps more importantly, provides interesting perspectives with respect to the modelling dynamic land allocation decisions based on the growing body of results related to dynamic discrete choices.

This section uses as much as possible the notations introduced in the preceding section. The farmer is assumed to have N plots of equal size ($n = 1,...,N$) and to decide for each which of the K crops to grow. It is first assumed the plots are "homogenous" in a sense to be defined later. According to the assumption of constant return to land, the profit associated with crop k on plot n is given by:

(12) $\pi_{kn}(p_k, \mathbf{w}) \equiv \pi_k(p_k, \mathbf{w}) + e_{kn}$.

The term $\pi_{kn}(p_k, \mathbf{w})$ is assumed to be known to the considered farmer. The term e_{kn} is unkown to the econometrician but the plots are assumed to be sufficiently homogenous for 1) $\pi_k(p_k, \mathbf{w})$ to be considered as the expected profit of growing crop k on any given plot and 2) considering that the term is identically and independently distributed (iid) across plots. The term e_{kn} is random from the econometrician's view point and, by definition, its mean is 0. It is further assumed 1) the e_{kn} terms are iid across plots and crops and 2) that they have a Weibull distribution minus e/a where e is the Euler constant and a is a scale parameter of the standard deviation which is provided by $a\sqrt{\pi/6}$.

According to these assumptions, the probability (as perceived by the econometrician) that the considered farmer chooses crop k for plot n is provided by:

$$(13)\quad \begin{aligned} P[k,n/\mathbf{p},\mathbf{w}] &= \frac{\exp\left[a^{-1}\pi_k(p_k,\mathbf{w})\right]}{\sum_{m=1}^{K}\exp\left[a^{-1}\pi_m(p_m,\mathbf{w})\right]} \\ &= \frac{\exp\left[a^{-1}\left(\pi_k(p_k,\mathbf{w})-\pi_K(p_K,\mathbf{w})\right)\right]}{1+\sum_{m=1}^{K-1}\exp\left[a^{-1}\left(\pi_m(p_m,\mathbf{w})-\right)\pi_K(p_K,\mathbf{w})\right]} \equiv P[k/\mathbf{p},\mathbf{w}] \end{aligned},$$

i.e., has a MNL form. It can be noted that for similar crops k and m, *e.g.* different cereals, the independence assumption of e_{kn} and e_{mn} may appear very restrictive. Similar crops may behave similarly on a given plant. If the crops can be allocated to mutually exclusive crop groups (nests) and the terms e_{kn} and e_{mn} of two crops of the same group are linked by the existence in those terms of an additive common factor being an extreme value random variable, then the probability of any crop k being chosen for plot n has a Nested MNL form.

It is important note that the parameters a used in this section have very different interpretation. The opposite remark applies concerning the motives for grouping the crops to form the nests of the Nested MNL crop shares.

The profit term $\pi_{kn}(p_k,\mathbf{w})$ is not associated with a fixed cost c_k similar to the one defined in section 2. This comes from the fact that this profit only considers the optimal allocations of crops to plots without considerations on the land allocation management costs. This would be the "ideal" choice of the farmer, *i.e.* his choice without constraints. Pursuing this logic, if the choice of crops are independent across plots, the expected (as perceived by the econometrician) share of plots allocated to crop k is given by $P[k/\mathbf{p},\mathbf{w}]$. Thus $P[k/\mathbf{p},\mathbf{w}]$ defines a measure of the acreage shares the farmer would choose if he was not constrained. In this sense $P[k/\mathbf{p},\mathbf{w}]$ is analogous to a stationary equilibrium choice of acreage share of crop k or a long term, *i.e.* with the possibility fixed factors quantity adjustments, choice of acreage share of crop k given prices $\mathbf{p}$ and $\mathbf{w}$. According to this logic and assuming that the farm is close to an equilibrium path, his dynamic choice of acreage shares can be approximated by a simple partial adjustment model of the log transformed acreage shares (see, e.g., Considine and Mount, 1984, for an example of this approach with MNL input cost shares). Denoting by $s_k(\mathbf{p}_t, \mathbf{w}_t)$ the share of land devoted to crop k in year t, the partial adjustment model is given by:

$$(14) \qquad \ln s_k(\mathbf{p}_t, \mathbf{w}_t) - \ln s_k(\mathbf{p}_{t-t}, \mathbf{w}_{t-1}) = r\left[\ln P\left[k/\mathbf{p}_t, \mathbf{w}_t\right] - \ln s_k(\mathbf{p}_{t-t}, \mathbf{w}_{t-1})\right] + \varepsilon_{kt}$$

where ε_{kt} is an approximation error term that includes the approximation error due to the use of the simple adjustment model as well as the error associated with the use of $P\left[k/\mathbf{p}, \mathbf{w}\right]$ in place of the true "target" choice of the farmer. Differentiation of equation (14) for crop k and crop K leads to the following estimating equation:

$$(15) \qquad \ln\left[\frac{s_k(\mathbf{p}_t, \mathbf{w}_t)}{s_K(\mathbf{p}_t, \mathbf{w}_t)}\right] = ra^{-1}\left[\pi_k(p_{k,t}, \mathbf{w}_t) - \pi_k(p_{k,t-1}, \mathbf{w}_{t-1})\right] + (1-r)\ln\left[\frac{s_k(\mathbf{p}_{t-1}, \mathbf{w}_{t-1})}{s_K(\mathbf{p}_{t-1}, \mathbf{w}_{t-1})}\right] + (\varepsilon_{kt} - \varepsilon_{Kt})\,.$$

Equation (15) is similar to equation (3). The r parameter defines the weight of the "target" or ideal choices with respect to the adjustment constraint. Its role is similar to that of parameter a in the structural MNL of section 2.

As it is defined as a long term or "target" acreage share, a natural generalisation of the model is to consider the choice of the farmer in period $t+1$, $t+2$, … Assuming for simplicity that the production dynamics is of order 1 and that farmer only considers $t+1$ with a discount factor d, results derived by Rust (1987) shows that the pay-off considered by the farmer in year t on any plot as a simple form. The production dynamics is represented by crop rotation effects. Let denote the profit of growing crop k on plot n where crop m was grown the preceding year by:

(16) $\pi_{kn/m}(p_{kt},\mathbf{w}_t)\equiv\pi_{k/m}(p_{kt},\mathbf{w}_t)+e_{knt}$

where the form of the $\pi_{k/m}(.)$ functions is known and the e_{knt} terms are iid across crops, plots and t and have the distribution defined above. In year t the e_{knt} terms are known to the farmer but the $e_{kn,t+1}$ terms are not. It is however that the farmer's perceived distribution of the $e_{kn,t+1}$ terms is the Weibull described above. According to this model, with prices known $(\mathbf{p}_{t+1},\mathbf{w}_{t+1})$ the farmer (and as well as the econometrician) knows that if he chooses crop k for plot n in t, the probability of his choosing crop ℓ in year $t+1$ on the same plot is given by:

(17a) $$P\left[\ell/k/\mathbf{p}_{t+1},\mathbf{w}_{t+1}\right]=\frac{\exp\left[a^{-1}\pi_{\ell/k}(p_{\ell,t+1},\mathbf{w}_{t+1})\right]}{\sum_{m=1}^{K}\exp\left[a^{-1}\pi_{m/k}(p_{m,t+1},\mathbf{w}_{t+1})\right]}$$

and thus his expected pay-off on plot n in year $t+1$ (as perceived in year t) is given by:

(17b) $$E\left[\pi_{/k}/\mathbf{p}_{t+1},\mathbf{w}_{t+1}\right]=a\ln\left[\sum_{\ell=1}^{K}\exp\left[a^{-1}\pi_{\ell/k}(p_{\ell,t+1},\mathbf{w}_{t+1})\right]\right],$$

i.e. the expected profit has the well-known log-sum form. Thus, in year t the (risk-neutral) farmer (who has a perfect foresight on $t+1$ prices) chooses the crop on plot n according to the pay-offs given by:

(17c) $$\pi_{k/m}(p_{kt},\mathbf{w}_t)+e_{knt}+da\ln\left[\sum_{\ell=1}^{K}\exp\left[a^{-1}\pi_{\ell/k}(p_{\ell,t+1},\mathbf{w}_{t+1})\right]\right].$$

For the econometrician, the probability that the farmer chooses k is provided by:

(17d) $$P\left[k/m/\mathbf{p}_t,\mathbf{w}_t,\mathbf{p}_{t+1},\mathbf{w}_{t+1}\right]=\frac{\exp\left[a^{-1}\left(\pi_{k/m}(p_{kt},\mathbf{w}_t)+da\ln\left[\sum_{\ell=1}^{K}\exp\left[a^{-1}\pi_{\ell/k}(p_{\ell,t+1},\mathbf{w}_{t+1})\right]\right]\right)\right]}{\sum_{q=1}^{K}\exp\left[a^{-1}\left(\pi_{q/m}(p_{qt},\mathbf{w}_t)+da\ln\left[\sum_{\ell=1}^{K}\exp\left[a^{-1}\pi_{\ell/q}(p_{\ell,t+1},\mathbf{w}_{t+1})\right]\right]\right)\right]}.$$

This term has the form of the choice probability of a nest in a Nested MNL discrete choice. The closed form of the expected pay-off in year $t+1$ allows further generalizations (see, *e.g.*, Rust, 1987) in particular in multiple periods optimization frameworks. Uncertainty about prices in $t+1$ can be handled using integration of the pay-off function (17b) according to the assumed distribution of prices $(\mathbf{p}_{t+1}, \mathbf{w}_{t+1})$ by simulation methods that are know widely used.

However simplifications arise in some cases. Crop *m* may provides the same effect on the profit of all succeeding crops, *e.g.* a given amount of nutrients. The application presented in the next section provides an example of such (crude) simplification.

4. Empirical illustrations

Data

This section presents two simple applications of the modelling frameworks presented in section 2 and 3, *i.e.* the structural MNL and discrete choices MNL frameworks for acreage shares. The application uses data from the French Farm Accountancy Data Network (FADN). It covers the 1988-2006 period and consists in a rotating (3 years *per* farm on average) panel of French arable crop producers for which winter wheat is the dominating product. The data set contains approximately 6000 observations. Years 1992 and 1993 were excluded due to the non recording of data on per hectare compensatory payments for the decreases in cereal and, oilseeds and protein crops prices due to the 1992 CAP reform. Only farms with records for two consecutive years were kept in the data set.

Typical recorded data, *e.g.* those made available by the FADN, provide for each farm i ($i=1,...,N$) and year t ($t=t_i,...,T_i$) detailed information on crop production. They provide the acreage (s_{kit}), yield (y_{kit}) and price at the farm gate (p_{kit}) for each crop k ($k=1,...,K$). But they only provide aggregate data on variable input (pesticides, fertilizers, seeds, …) expenditures whereas input price indices are made available at the regional level. The different variable inputs (fertilizers, pesticides, energy, seeds, …) were aggregated into a variable input for simplicity. It can be noted that results of other versions of our models that the input use reductions are mainly due to reduction in fertilizer uses. However, we had to handle the problem of the allocation of this variable input to the different crops. The approach we used surely is a weakness of our empirical illustrations as will be discussed later. X_{it} denotes the quantity of input k ($k=1,...,K$) purchased by farm i during year t and w_{ti} denotes the corresponding price index. These data also provide measures of the quantities of inputs which are (quasi-)fixed in the short-run, *i.e.* the total quantity of land. All quantities are defined in € of 2000.

The specified models consider six crops: wheat, other cereals (barley and corn), oilseeds (mainly rapeseed) and protein crops (mainly peas), sugar beets, potatoes and miscellaneous crops, and fodder crops. Sugar beet acreages were considered as exogenous due to the quota system implement in the UE. Acreages of potatoes and miscellaneous crops were also considered as a exogenous since most of their output is subject to contracts and because their acreage is rather limited. Fodder crop acreage (mainly silage corn) was also considered as exogenous due to feeding constraints. However, the effects of fodder crop acreages were on the acreage of other cereals were considered: farmers change grain corn into silage corn when silage corn production is expected to be low due to climatic conditions. This occurred in 2003 and 2005 due to dramatic droughts.

Theoretical models

The quadratic functional form is chosen for the yield functions for two reasons. First its associated dual functions have simple functional forms. The unavailability of data on input uses per crop requires the specification of (dual) input demand functions *per* crop. Second, the quadratic production function can be parameterized in a form which is fairly easy to interpret and which allows to consider error terms in a simple way. The yield equations are defined as:

(18a) $y_k = \beta_k - \frac{1}{2}\gamma_k \ \ (\delta_k - x_k)^2$ with $\beta_k > 0,\ \gamma_k > 0$ and $\delta_k > 0$

where x_k is the quantity of variable input used *per* hectare devoted to crop *k*. In this primal framework, the β_k and δ_k parameters have direct interpretations: δ_k is the vector of variable input quantities required to achieve the maximum yield β_k. The term γ_k determines the curvature of the yield function and, as a result, greatly determines the price effects. These parameters have direct "agronomic" interpretations that permitted us to "check' our results with agricultural scientists. The maximisation of the gross margin for crop *k* leads to the following *per* hectare supply function :

(18b) $y_k(p_k, w) = \beta_k - \frac{1}{2}\gamma_k (w/p_k)^2$,

variable input demand function :

(18c) $x_k(p_k, w) = \delta_k - \gamma_k (w/p_k)$

and gross margin functions :

(18d) $\pi_k\left(p_k, w\right) = p_k \left[\beta_k - \delta_k \ \ \frac{w}{p_k} + \frac{1}{2}\gamma_k (w/p_k)^2\right] = \left[p_k\beta_k - w\delta_k + p_k \frac{1}{2}\gamma_k (w/p_k)^2\right].$

These functions have simple functional forms suitable for empirical use.

Two versions of the theoretical model of crop acreages were used. For the structural MNL framework model we use the following acreage share equations:

(19a) $\ln(s_k/s_1) = a^{-1}\left[\pi_k(p_k, w) - \pi_1(p_1, w)\right] + \ln d_k - \ln d_1$ for $k > 1$.

Wheat was used as the benchmark crop (1). Assuming that the reference acreage shares d_k were of the standard MNL form, the differences in log "reference" acreage shares $\ln d_k - \ln d_1$ were defined as a linear function of the characteristics (**c**) of the farm that are presented in the next paragraph. The model was reparametrized to facilitate the interpretation of the a^{-1} parameter. The equation estimated has the following form:

$$\text{(19c)} \quad \ln(s_k/s_1) = \alpha\left[\pi_k(p_k, w) - \pi_1(p_1, w) - c_{k1}\right] + (1-\alpha)(\zeta_{0k} + \boldsymbol{\zeta}_k'\mathbf{c}),$$

the parameter α denoting the weight of the economic factors in the acreage choices decisions.

For the discrete MNL framework model we used the following simple acreage share equations:

$$\text{(20)} \quad \ln(s_{k,t}/s_{1,t}) = ra^{-1}\left[\pi_k(p_k, w) - \pi_1(p_1, w)\right] + (1-r)\ln(s_{k,t-1}/s_{1,t-1}) + \varepsilon_t.$$

Econometric models

In what follows, the indices 1, 2 and 3 respectively denote wheat, other cereals and, oilseeds ans protein crops.

The difficulty for the specification of the econometric models in this context are threefold. It is necessary to take into account the heterogeneity of the production conditions faced by the farms. This heterogeneity mainly reflects the impacts of pedo-climatic conditions. In order to account for the heterogeneity of soil quality and of local standard climatic conditions, a "quality index" was created. For each farm *i* obsverved in period *t* and $t-1$, this index was defined as:

$$\text{(21)} \quad q_i \equiv \frac{y_{1,i,t-1} - y_{1,Med,t-1}}{y_{1,Max,t-1} - y_{1,Min,t-1}}.$$

The $y_{1,Med,t-1}$, $y_{1,Med,t-1}$ and $y_{1,Min,t-1}$ terms respectively denote the median, 99% quantile and 1% quantile of the yield of wheat in the sample in year $t-1$. This index lies between 0 and 1. It is defined on wheat yield to the specialisation of the farms of the sample. It is defined on a year by year basis due to data availability and to control for general year specific conditions, this also explain why it was defined or supposed to only depends on *i*). All computations (including estimations) were weighted to correct the sample content with respect to year/*région* stratification. The q_i index is also divided by the $y_{1,Max,t-1} - y_{1,Min,t-1}$ to account for differences in the amplitude in yields due to specific climatic conditions. It was defined as a measure of the rank of the conditions faced by the farms. This index has severe drawbacks as a measure of farm specific production conditions since it also depends of the farmers' decisions. However, for the specific purpose of this study, these drawbacks can also be seen as an advantage due to the difficulties associated with the input allocation to crops. Trends were introduced in the yield functions to account for (embodied in seeds) technical changes and general climatic conditions.

The second difficulty comes from the introduction of the structural error terms, and especially the part of the error terms that represent elements known to the farmers but unknown to the econometrician. These error terms determine the moment conditions that can be used to identify the parameters of the model. In this context the q_i index can be seen as a variable closely linked to the so-called individual fixed effects in panel data econometrics. As shown by Carpentier and Weaver (1996, 1997) these fixed effects are important to control for the heterogeneity in production conditions and in choices of farmers to achieve high yields or not, especially for te French grain crop sector. It is important to note that this index has either the expected effects or easily interpretable effects in the estimated models. The results of this study in this respect are consistent with those of Carpentier and Weaver. The third main difficulty associated with the specifications of the econometrics models is the representation of the crop rotations effects on yields and inputs uses. Crude approximations were used in this study. They mainly focus on the beneficial effects (important) of sugar beets and potatoes as predecessors of wheat and of oilseeds (moderated) and protein crops (important) on wheat. Indeed, the rotations in France are organized to benefit to wheat. Corn easily support monoculture, barley usually follows wheat in standard rotations excepted where it is subject to contracts with breweries. As argued in section 2 and 3, dynamic optimization aspects are likely to be a consequence of the crop rotation effects. A very simple test conducted in the discrete MNL model tends to confirm this hypothesis.

The basic econometric specification for the yield functions is defined by:

(22a) $$y_{kit} = \beta_{kit} - \frac{1}{2}\gamma_{ki}\,(w_{it}/p_{kit})^2\,,\quad k = 1,...,3$$

where:

(22b) $$\beta_{kit} = \beta_{k0} + \beta_{k1}q_i + \beta_{k2}q_i^{\,2} + \beta_{k3}(t-1989) + \beta_{k4}(t-1989)^2 + u_{kit}$$

and:

(22c) $$\gamma_{kit} = \gamma_{k0} + \gamma_{k1}q_i + \gamma_{k2}q_i^{\,2}\,.$$

This basic model accounts for heterogeneity in the parameters of the yield functions as quadratic functions of the q_i index and considers a quadratic trend in the "maximum yield" parameter β_{kit}. The structural error term is also introduced β_{kit} leading a standard dual yield function specification. The basic econometric specification for the input demand function is rather different since it considers the total use of variable input (*per* hectare)

(23a) $$X_{it} = \sum_{k=1}^{3} s_{kit}\left[\delta_{ki} - \gamma_{ki}\quad (w_{it}/p_{kit})\right] + \sum_{k=4}^{5} s_{kit}\eta_{kit} + v_{it}$$

where:

(23b) $$\delta_{kit} = \delta_{k0} + \delta_{k1}q_i + \delta_{k2}q_i^{\,2} + e_{kit}\,,\quad k = 1,2,3$$

and:

(23b) $$\upsilon_{kit} = \upsilon_{k0} + \upsilon_{k1}q_i + \upsilon_{k2}q_i^{\,2} + \upsilon_{k3}(t-1989) + \upsilon_{k4}(t-1989)^2 + \upsilon_{k5}(t-1989)^3 + \varsigma_{kit}\,,\quad k = 4,5,6\,.$$

Other characteristics of farms (total area, …) were introduced in some terms. In particular the share of irrigated land where introduced in some cases. The increase of irrigated corn production explains a large part of its yield increase. The "maximum" yield and input parameters of the yield function were completed by the effects of the acreage share of sugar beets and potatoes to account for beneficial rotation effects leading to:

(24a) $\beta_{kit} = \beta_{k0} + \beta_{k1}q_i + \beta_{k2}q_i^2 + \beta_{k3}(t-1989) + \beta_{k4}(t-1989)^2 + \beta_{k5}s_{4=bp,i,t-1}$

and:

(24b) $\delta_{kit} = \delta_{k0} + \delta_{k1}q_i + \delta_{k2}q_i^2 + \delta_{k2}s_{4=bp,i,t-1} + e_{kit}$.

The acreage share equations of the structural MNL model are given by :

(25a) $\ln(s_{kit}/s_{1it}) = \alpha[\pi_{kit} - \pi_{1it} - c_{k10} + dp_{kit} - dp_{1it}] + (1-\alpha)(\zeta_{0k} + \boldsymbol{\zeta}_k'\mathbf{c}_{kit}) + \theta_{kit}, \quad k = 2,3$

where:

(25b) $\pi_{kit} = p_{kit}\beta_{kit} - w_{kit}(\delta_{kit} - e_{kit}) + p_{kit}\frac{1}{2}\gamma_{ki}(w_{it}/p_{kit})^2, \quad k = 1,2,3$ a

and where the $\boldsymbol{\zeta}_k'\mathbf{c}_{kit}$ are defined as the sum of linear and quadratic effects of q_i, the current acreage share of sugar beets and potatoes, the current share of fodder crops and the past acreage share of wheat. The terms dp_{kit} are the direct payments perceived form the EU for each hectare of a given crop. These payments where implement to compensate farmers for the decrease in price supports provided by the CAP. They are high for oilseeds and protein crops but decrease for all crops progressively. In 2006, the area planted in grain crops were eligible for the same direct payment.

It is important to note that e_{kit} terms do no appear in the profit level used in the acreage equations. This assumption is crucial for the consistency of the input demand equation since it implies that the error term of this equation does not contain terms of the form $e_{kit}s_{kit}$. This terms would imply that the $e_{kit}s_{kit}$ terms are endogenous in the equation. The expectation of these terms would not even be null. The validity of this assumption relies on the assumption that the terms q_i all specific effects of farmer i that explains its acreage decisions. In this context, e_{kit} can be assumed to be a term that is only known to the farmer after his acreage choices. This assumption is admittedly crucial and subject to caution. Another way to handle this problem would be choose a distribution for the vector of terms constructed with the q_i to integrate all relevant expectations by use of simulation techniques.

The acreage share equations of the discrete choice MNL model are simply given by:

(26) $$\ln(s_{ki,t}/s_{1i,t}) = ra^{-1}\left[\pi_{kit} - \pi_{1it} - c_{k1} + dp_{kit} - dp_{1it}\right] + (1-r)\ln(s_{ki,t-1}/s_{1i,t-1}) + \varepsilon_{kit}, \quad k = 2,3.$$

The estimators of parameters of the models consisting in system with three yield equations, one input demand equation and two acreage shares equations were constructed within the Generalized Method of Moment framework. These estimators were constructed based on the orthogonality conditions defined the cross product of the error term of each of the 6 equation of the model with each or their explanatory variables on a equation *per* equation basis. Indeed the vast majority of these explanatory variables are exogenous in all equations of the two systems. The acreage shares equations of wheat, of other cereals and of oilseeds and protein crops are exogenous in the input demand equation but not in the others. The GMM estimator was used in its version robust to heteroskedasticity of unknown form and does not exclude correlation of error terms across equations (see, *e.g.*, Wooldridge, 2003).

Main results

Both models yield similar results with respect to the input demand and yield function parameters Tables 1a and 1b. The fit of the data to the model is rather good given the nature of the data. The yield function of wheat is best explained by the data. This also illustrates the fact the wheat production is less risky than that of the other crops. The parameters associated to the index q_i have the expected signs according : they have positive effects on the maximum yield and maximum input parameters. The impacts of this index are important in the results as shown later. The estimated trends for the maximum yield parameters of wheat and other cereals are increasing during the 90's but start to decrease in the 00's. This main partially explained by the facts 2001 was particularly humid whereas 2003, 2005 and 2006 were particularly dry. The trend on the yield of oilseeds and of protein crops is decreasing. This essentially explained by the fact that peas, the main protein crop, is sensible to a soil disease for which no pesticide exists. The acreage of this crop decrease but rather slowly, thanks to direct payments (per hectare planted) provided by the EU.

The price parameters γ depends on q_i and implies a own-price elasticity of input uses of about .3 for wheat, almost null for the other cereals and of about .5 for oilseeds and protein crops. These results are rather standard for wheat and oilseeds and protein crops. The non-responsiveness of the input demand for the other cereals is more surprising. Other results tend to show that the input demand for corn is more price elastic than that of barley. This may be explained by the contracts for barley. Tables 3a and 3b show that the drop in prices decreased the input demand for wheat by 22% and the input demand for oilseeds and protein crops by 36%. Estimations of these models for specific inputs show that the decrease in the demand of the aggregate variable input is mainly due to a decrease in fertilizer uses.

Table 4a and 4b show interesting results with respect to the heterogeneity of price effects. The price responsiveness of the production practices of oilseeds and protein crops is more important in *régions* with high wheat yields: Nord-Pas-de-Calais, Picardie, Champagne and Ile-de-France. These *régions* are also characterized by large sugar beets and potatoes acreages. This may be interpreted by the fact that oilseeds and protein crops are mainly used as preceding crops for wheat in these *régions* where they complete the role of sugar beets and potatoes in this respects. These crops were also profitable in the 80's and in the early 90, but the price decrease that occurred in 1992 (-50%) seems to have eliminated the economic incentives to put much effort on these crops with low returns. The story is different in *régions* with low agricultural potential , *e.g.* Pyrénées and Poitou-Charentes, where oilseeds' yields are closer to those of wheat than they are in high potential *régions*. It can be noted that the effects of sugar beets-wheat rotation on wheat yield and on input uses for wheat are highly significant.

The estimates of the acreage equations are more contrasted. Table 3a presents the main results for the structural MNL model. The effects of the index q_i were expected: in high wheat yield areas farmers devote more land to wheat. The effects of the shares of the different crops shares on the acreages of other cereals and of oilseeds and protein crops were expected. A high acreage of sugar beets and potatoes benefit to wheat acreages, the crop that benefits the more from this rotation (with barley). Where farmers produce more fodder (mainly corn) they devote less land to the other cereals (including corn) but they devote more land to oilseeds to provide rotations possibilities for their cereals. When farmers grow more wheat, they also produce less of the other cereals. The estimate of the α parameter indicate that, according, to the results of the models the short run profits and direct payments only explain 4.5% of the part determined by the model, i.e.

58% of the acreage variation for the cereals and 44% of the acreage variation for oilseeds and protein crops. This result is rather disappointing. It shows the specified model does not allow to explain much of the variation of crop acreage across time and regions. This certainly calls for improvements of the specification of crop rotation effects (and dynamic optimization aspects). But this may also comes from the political context. Price supports to grain crops sharply declined in the EU during the 90's. But this decrease in prices was compensated by direct payments computed to avoid large decreases in producers' revenue. This direct payments were computed grain crop per grain crop at the *département* level (France is divided into 95 *départements*). This tended to "fix" the grain crop acreages.

The estimates related to the discrete choice MNL model are presented in table 3b. They are more promising. The partial adjustment model presents a good fit to the data with R2 criterion equal to .66 and .70. This is not surprising for a model explaining acreages in year t by acreages in year $t-1$. However, the estimate of the r parameter shows that the short run profits and direct payments determine 23.5% of the explanatory power of the model. This means that acreages depends on short run economic incentives. This is also due to the fact that the global (direct and through price) support to grain crop producers' revenue tends to decreases. It is also interesting to note that the "fixed costs" parameters tend to show that the model predict underestimate the acreages of oilseeds and other cereals. It is interested to note that a modification model incorporating the effects of the oilseeds/protein crops-wheat rotation along with a crude measure of the anticipation of the economic effects on wheat profit provided a correction for the underestimation problem: the considered rotation has limited effects on wheat yields but significantly decreases inputs uses for wheat. Anticipation of these effects increases the profitability of these crops.

5. Concluding remarks

Two approaches are presented provide theoretical justifications for using Logit acreage share models: the Logit shares can be derived from a well defined profit function or derived as the result of a set of discrete choices. Of course, the theoretical background presented in this paper can be further developed.

Both theoretical frameworks allow to define generalizations of the standard Logit shares. These generalizations build on developments of the MNL Logit framework for modelling discrete choices and seek to define models that are flexible and empirically tractable.

Of these, potential developments the extension of MNL framework to account for crop rotation effects in production and dynamic optimization by farmers. This is confirmed by the applications presented in the paper. The discrete choice MNL framework seems more flexible and thus more easy to generalize. These generalizations could benefit from the rapidly expanding literature on dynamic discrete choices.

The main draw back of the MNL framework is that it rules automatically out corner solutions. This certainly calls for original approaches for corner solution modelling.

Table 1a. Estimates of the yield functions and demand function parameters, Structural MNL model

	Price effects parameters			Yield functions parameters						Demand function parameters				R^2
	γ_0	γ_1	γ_2	β_0	β_1	β_2	β_3	β_4	β_5	δ_0	δ_1	δ_2	δ_3	
Wheat	0.992	0.540	-1.346	7.478	3.263	-0.601	0.095	-0.003	1.479	5.182	1.240	-2.150	-3.238	0.44
	(0.131)	(0.221)	(0.497)	(0.051)	(0.102)	(0.229)	(0.010)	(0.000)	(0.143)	(0.161)	(0.271)	(0.574)	(0.703)	
Other cereals	-0.288	-0.631	-1.858	8.025	2.308	-1.058	0.105	-0.003		4.149	-0.077	-1.470		0.33
	(0.224)	(0.242)	(0.531)	(0.111)	(0.118)	(0.263)	(0.012)	(0.000)		(0.257)	(0.310)	(0.594)		
Oilseeds and protein crops	1.269	1.500	0.101	5.519	2.661	-0.128	0.140	-0.006		4.911	3.130	0.689		0.36
	(0.160)	(0.319)	(0.647)	(0.065)	(0.141)	(0.283)	(0.012)	(0.001)		(0.220)	(0.393)	(0.750)		
Total input demand														0.46

(Estimated std errors of the estimators are in parentheses)

Table 1b. Estimates of the yield functions and demand function parameters, Discrete choice MNL model

	Price effects parameters			Yield functions parameters						Demand function parameters				R^2
	γ_0	γ_1	γ_2	β_0	β_1	β_2	β_3	β_4	β_5	δ_0	δ_1	δ_2	δ_3	
Wheat	1.009	0.488	-1.348	7.472	3.249	-0.628	0.099	-0.003	1.546	5.024	1.005	-1.504	-5.862	0.44
	(0.121)	(0.221)	(0.497)	(0.050)	(0.102)	(0.229)	(0.010)	(0.000)	(0.142)	(0.129)	(0.232)	(0.516)	(0.585)	
Other cereals	0.315	-0.757	-1.787	7.770	2.184	-0.849	0.096	-0.003		4.856	0.356	-1.207		0.33
	(0.140)	(0.241)	(0.530)	(0.074)	(0.117)	(0.262)	(0.011)	(0.001)		(0.170)	(0.272)	(0.555)		
Oilseeds and protein crops	1.948	3.139	0.403	5.842	3.361	-0.009	0.127	-0.006		5.806	4.235	0.196		0.35
	(0.133)	(0.272)	(0.644)	(0.054)	(0.123)	(0.281)	(0.011)	(0.001)		(0.166)	(0.305)	(0.665)		
Total input demand														0.45

(Estimated std errors of the estimators are in parentheses)

Table 2a. Estimates of the acreage decision parameters, Structural MNL model

	Other cereals / wheat		Oilseeds and protein crops / wheat	
ζ_0	3.479	(0.078)	1.215	(0.063)
ζ_1	-0.155	(0.031)	0.006	(0.036)
ζ_2	0.090	(0.064)	0.276	(0.068)
$\zeta_{fo\,1}$	-2.786	(0.478)	1.641	(0.457)
$\zeta_{fo\,2}$	-3.427	(3.088)	-12.925	(2.937)
$\zeta_{sb\,1}$	-3.491	(0.369)	-3.924	(0.220)
$\zeta_{sb\,2}$	-1.882	(0.591)	1.741	(0.778)
$\zeta_{wh\,1}$	-7.509	(0.309)	-5.937	(0.291)
$\zeta_{wh\,2}$	1.971	(0.387)	4.204	(0.363)
R²	0.58		0.44	
α	0.045 (0.004)			

(Estimated std errors of the estimators are in parentheses)

Table 2b. Estimates of the acreage decision parameters, Discrete choice MNL model

	Other cereals / wheat		Oilseeds and protein crops / wheat	
Fixed costs	-1.003	(0.153)	-0.857	(0.176)
R²	0.70		0.64	
a	0.235 (0.006)			
r	3.703 (0.173)			

(Estimated std errors of the estimators are in parentheses)

Table 3a. Average price ratio, yield and estimated input uses per crop and year (in 100 € 2000), Structural MNL model

	Wheat			Other cereals			Oilseeds and protein crops		
Year	**w/p**	**Yield**	**Input use**	**w/p**	**Yield**	**Input use**	**w/p**	**Yield**	**Input use**
1989	1.76	7.39	3.94	1.72	7.17	4.45	2.06	6.34	4.57
1990	1.77	7.59	3.90	1.69	6.83	4.48	2.16	5.82	4.52
1991	1.69	7.88	3.83	1.71	7.39	4.47	1.95	6.09	4.41
1994	1.39	7.46	3.76	1.26	7.11	4.50	1.09	5.67	3.69
1995	1.26	7.10	3.61	1.24	6.82	4.48	1.03	5.62	3.39
1996	1.25	7.79	3.55	1.31	7.48	4.50	0.97	5.95	3.27
1997	1.26	7.52	3.59	1.24	7.67	4.52	1.10	6.50	3.69
1998	1.18	8.68	3.46	1.12	8.00	4.47	1.15	6.44	3.61
1999	1.08	8.24	3.39	1.03	8.07	4.46	1.06	6.40	3.45
2000	1.04	8.11	3.34	1.02	7.97	4.46	0.89	5.53	3.08
2001	1.00	7.48	3.30	1.00	7.35	4.44	1.00	5.11	3.24
2002	1.06	8.33	3.34	1.00	8.03	4.44	1.17	5.95	3.62
2003	0.98	6.97	3.25	0.93	6.84	4.42	1.14	5.63	3.54
2004	1.05	8.57	3.29	1.06	8.24	4.47	1.15	6.32	3.58
2005	1.02	7.81	3.20	0.91	7.84	4.38	1.10	6.03	3.28
2006	0.91	7.61	3.06	0.90	7.49	4.36	1.04	5.29	3.21

Table 3b. Average price ratio, yield and estimated input uses per crop and year (in 100 € 2000), Discrete choice MNL model

	Wheat			Other cereals			Oilseeds and protein crops		
Year	**w/p**	**Yield**	**Input use**	**w/p**	**Yield**	**Input use**	**w/p**	**Yield**	**Input use**
1989	1.76	7.39	3.83	1.72	7.17	4.32	2.06	6.34	4.63
1990	1.77	7.59	3.83	1.69	6.83	4.34	2.16	5.82	4.55
1991	1.69	7.82	3.76	1.71	7.39	4.31	1.95	6.09	4.42
1994	1.39	7.46	3.69	1.26	7.11	4.36	1.09	5.67	3.82
1995	1.26	7.10	3.57	1.24	6.82	4.33	1.03	5.62	3.49
1996	1.25	7.79	3.51	1.31	7.48	4.37	0.97	5.95	3.36
1997	1.26	7.52	3.57	1.24	7.67	4.42	1.10	6.50	3.71
1998	1.18	8.68	3.47	1.12	8.00	4.33	1.15	6.44	3.63
1999	1.08	8.24	3.39	1.03	8.07	4.34	1.06	6.40	3.45
2000	1.04	8.11	3.34	1.02	7.97	4.35	0.89	5.53	3.06
2001	1.00	7.48	3.31	1.00	7.35	4.31	1.00	5.11	3.23
2002	1.06	8.33	3.35	1.00	8.03	4.33	1.17	5.95	3.58
2003	0.98	6.97	3.24	0.93	6.84	4.31	1.14	5.63	3.50
2004	1.05	8.57	3.29	1.06	8.24	4.38	1.15	6.32	3.52
2005	1.02	7.81	3.19	0.91	7.84	4.27	1.10	6.03	3.21
2006	0.91	7.61	3.04	0.90	7.49	4.26	1.04	5.29	3.11

Table 4a. Average yield, estimated price parameter and acreage choices per crop and *région* (in 100 € 2000), Structural MNL model

	Wheat			Other cereals			Oilseeds and protein crops			Sugar beets
Région	**Yield**	γ	s	**Yield**	γ	s	**Yield**	γ	s	s
Nord-Pas-de-Calais	8.96	0.96	0.43	8.18	0.06	0.19	6.61	3.52	0.12	0.14
Picardie	8.66	0.99	0.46	7.95	0.17	0.17	6.63	3.10	0.18	0.11
Champagne-Ardenne	8.42	0.95	0.36	7.64	0.21	0.26	6.32	2.43	0.23	0.07
Ile-de-France	8.02	0.95	0.47	7.54	0.15	0.24	6.09	2.36	0.23	0.05
Haute-Normandie	7.63	0.90	0.48	6.97	0.37	0.17	5.79	2.33	0.28	0.02
Centre	7.38	0.86	0.44	7.55	0.27	0.27	5.67	1.69	0.26	0.01
Franche-Comté	7.26	0.83	0.35	7.83	0.07	0.29	5.59	1.05	0.32	0.00
Bourgogne	6.97	0.78	0.41	6.61	0.37	0.25	5.39	1.22	0.30	0.01
Poitou-Charentes	6.47	0.66	0.33	7.80	0.10	0.33	4.87	1.07	0.31	0.00
Midi-Pyrénées	5.70	0.40	0.24	6.43	0.16	0.35	4.36	0.67	0.38	0.00

Table 4b. Average yield, estimated price parameter and acreage choices per crop and *région*, Discrete choice MNL model

	Wheat			Other cereals			Oilseeds and protein crops			Sugar beets
Région	**Yield**	γ	s	**Yield**	γ	s	**Yield**	γ	s	s
Nord-Pas-de-Calais	8.96	0.96	0.43	8.18	-0.11	0.19	6.61	2.80	0.12	0.14
Picardie	8.66	0.99	0.46	7.95	0.05	0.17	6.63	2.52	0.18	0.11
Champagne-Ardenne	8.42	0.96	0.36	7.64	0.07	0.26	6.32	2.38	0.23	0.07
Ile-de-France	8.02	0.96	0.47	7.54	0.18	0.24	6.09	2.15	0.23	0.05
Haute-Normandie	7.63	0.92	0.48	6.97	0.23	0.17	5.79	1.89	0.28	0.02
Centre	7.38	0.88	0.44	7.55	0.25	0.27	5.67	1.75	0.26	0.01
Franche-Comté	7.26	0.86	0.35	7.83	0.32	0.29	5.59	1.51	0.32	0.00
Bourgogne	6.97	0.81	0.41	6.61	0.28	0.25	5.39	1.46	0.30	0.01
Poitou-Charentes	6.47	0.69	0.33	7.80	0.29	0.33	4.87	1.13	0.31	0.00
Midi-Pyrénées	5.70	0.44	0.24	6.43	0.18	0.35	4.36	0.66	0.38	0.00

References

Anderson, S. P., A. de Palma, and J. Thisse. 1992. *Discrete Choice Theory of Product Differentiation.* Cambridge, Mass. MIT Press.

Berry, S. T. 1994. "Estimating Discrete-Choice Models of Product Differentiation." *RAND J. Econ.* 25:242-262.

Carpentier, A., and R. D. Weaver. 1996. "Intertemporal and Interfirm Heterogeneity : Implications for Pesticide Productivity." *Can. J. Agr. Econ*. 44:219-236.

Carpentier, A., and R. D. Weaver. 1997. "Damage Control Productivity : An Input Abatement Approach."*Amer. J. Agr. Econ.* 52:11-22.

Caswell, M., and D. Zilberman. 1985. "The Choice of Irrigation Technologies in California." *Amer. J. Agr. Econ.* 67:224-234.

Chavas, J.-P. and T. Holt. 1990. "Acreage Decisions under Risk : The Case of Corn and Soybeans." *Amer. J. Agr. Econ.* 72:529-538.

Chambers, R., and R. Just. 1989. "Estimating Multiouput Technologies." *Amer. J. Agr. Econ.* 71:980-95.

Considine, T. J., and T. D. Mount. 1984. "The Use of Linear Logit Models for Dynamic Input Demand Systems." *Rev. Econ. and Statist.* 66:434-443.

Coyle, B. T. 1992. "Risk Aversion and Price Risk in Duality Models of Production : A Linear Mean-Variance Approach." *Amer. J. Agr. Econ.* 74:849-859.

Eckstein, Z. 1984. "A Rational Expectations Model of Agricultural Supply." *J. of Pol. Econ.* 92:1-19.

Henessy, D. A. 2007. "On Monoculture and the Structure of Crop Rotations." *Amer. J. Agr. Econ.* 88:900-914.

Howitt, E. 1995. "Positive Mathematical Programming." *Amer. J. Agr. Econ.* 77:329-342.

Jones, C. T. 1995. “A Dynamic Analysis of Interfuel Substitution in U.S. Industrial Energy Demand.” *J. Bus. Econ. Stat.* 13: 459-65.

Just, R. E., Zilberman, D., Hochman, E., and Z. Bar-Shira. 1990. “Input Allocation in Multicrop Systems.” *Amer. J. Agr. Econ.* 72:200-209.

Just, R. E., Zilberman, D., and E. Hochman. 1983. „Estimation of Multicrop Production Functions.” *Amer. J. Agr. Econ.* 65:770-780.

Lichtenberg, E. 1989. “Land Quality, Irrigation Development and Cropping Pattern in the Northern High Plains.” *Amer. J. Agr. Econ.* 71:187-94.

Moore, M. R., and D. H. Negri, D. H. 1992. “A Multicrop Production Model of Irrigated Agriculture, Applied to Water Allocation Policy of the Bureau of Reclamation.” *J. Agr. Res. Econ.* 17:29-43.

Ozarem, P. F., and J.A. Miranowsky. 1994. “A Dynamic Model of Acreage Allocation with General and Crop-Specific Soil Capital.” *Amer. J. Agr. Econ.* 76:385-395.

Rust, J. 1987. “Optimal Replacement of GMC Bus Engines: An Empirical Model of Harold Zurcher.” *Econometrica* 55:999-1033.

Thomas, A. 2003. “A Dynamic Model of On-Farm Integrated Nitrogen Management.” *Eur. Rev. Agr. Econ.* 30:439-460.

Train, K. 2003. *Discrete Choice Methods with Simulation*. Cambridge University Press.

Verboven, F. 1996. “The Nested Logit Model and Representative Consumer Theory.” *Economics Letters* 50:335-354.

Wooldridge, J.M., 2002. *The Econometrics of Cross-Sections and Panel Data.* The MIT press.

Wu, J., and K. Segerson. 1995. “The Impact of Policies and Land Characteristics on Potential Groundwater Pollution in Wisconsin.” *Amer. J. Agr. Econ.* 77:1033-47.