



The World's Largest Open Access Agricultural & Applied Economics Digital Library

This document is discoverable and free to researchers across the globe due to the work of AgEcon Search.

Help ensure our sustainability.

Give to AgEcon Search

AgEcon Search

<http://ageconsearch.umn.edu>

aesearch@umn.edu

*Papers downloaded from **AgEcon Search** may be used for non-commercial purposes and personal study only. No other use, including posting to another Internet site, is permitted without permission from the copyright owner (not AgEcon Search), or as allowed under the provisions of Fair Use, U.S. Copyright Act, Title 17 U.S.C.*

No endorsement of AgEcon Search or its fundraising activities by the author(s) of the following work or their employer(s) is intended or implied.



Using farm accountancy data to estimate crop rotation effects

F. Femenia¹; A. Carpentier¹; A. Gohin¹; R. Soudjadin²

1: INRA, SMART-LERECO, France, 2: INRA, SLART-LERECO, France

Corresponding author email: fabienne.femenia@inra.fr

Abstract:

Crop rotations effects are key features of environmentally friendly crop production practices but they are poorly documented. The aim of this article is to present an estimation approach of these effects based on farm accountancy data. Estimating crop rotation effects on yield and variable input use levels from farm accountancy data is a challenging issue since farmers' crop sequence choices are not observed. We propose to devise this estimation problem as a Bi-level Programming (BLP) problem designed to estimate the crop rotation effects while simultaneously reconstructing farmers' unobserved crop sequences from farmers' observed (current and previous) crop acreages. Our estimation approach is based on a well-defined statistical background. It relies on simple crop sequence yield and input use models as well as on an assumption stating that farmers are economically rational when deciding their crop sequence acreages. Our estimation approach also makes use of expert knowledge information on crop rotation effects for 'guiding' the construction of the crop rotation effect estimators. An illustrative application based on French farm accountancy data demonstrate that the proposed estimation approach yields meaningful crop rotation estimates when suitably implemented.

Acknowledgment: This research has benefitted from financial support from the European Commission.

JEL Codes: Q01, Q18

#1160



Using farm accountancy data to estimate crop rotation effects

Abstract

Crop rotations effects are key features of environmentally friendly crop production practices but they are poorly documented. The aim of this article is to present an estimation approach of these effects based on farm accountancy data. Estimating crop rotation effects on yield and variable input use levels from farm accountancy data is a challenging issue since farmers' crop sequence choices are not observed. We propose to devise this estimation problem as a Bi-level Programming (BLP) problem designed to estimate the crop rotation effects while simultaneously reconstructing farmers' unobserved crop sequences from farmers' observed (current and previous) crop acreages. Our estimation approach is based on a well-defined statistical background. It relies on simple crop sequence yield and input use models as well as on an assumption stating that farmers are economically rational when deciding their crop sequence acreages. Our estimation approach also makes use of expert knowledge information on crop rotation effects for 'guiding' the construction of the crop rotation effect estimators. An illustrative application based on French farm accountancy data demonstrate that the proposed estimation approach yields meaningful crop rotation estimates when suitably implemented.

Introduction

Agri-environmental policies are now assessed based on environmental criteria and are increasingly designed by referring to environmentally friendly crop production practices. Yet crop rotations are key features of these practices and should thus be taken into account in economic models aimed at evaluating these policies.

Adapting economic model of farmers' production decisions to account for crop rotations is a challenging issue, due to information lacking on crop rotation effects. In fact, the use and the effects of crop rotations are poorly documented (see, *e.g.*, Meynard *et al* (2013)). Farmers' crop sequence acreages are in fact rarely recorded, at least in large datasets suitable for statistical analyses, and crop rotation effects on yields and input uses are measured mostly based on experimental data and only for a few major crop pairs.

Econometric models aimed at accounting for farmers' using crop rotation effects were developed based on the dynamic acreage choice model proposed by Eckstein (1984) and have been employed since then (see, *e.g.*, Tegene *et al* (1989), Ozarem and Miranowski (1994), Thomas (2013), Vitale *et al* (2013)). These models allow estimating crop rotation effects without observing farmers' crop choice and their estimation only requires standard farm accountancy data. Nevertheless, as shown by Carpentier and Gohin (2015), they rely on very crude modelling frameworks as regards to crop rotation effects.

Theoretical models of farmers' dynamic acreage choices have been substantially improved over time and provide a useful basis to propose extensions of simulation models to dynamic settings accounting for crop rotations. Carpentier and Gohin (2015), notably, propose a model based on a dynamic programming framework aimed at computing optimal stationary crop sequence choices. This modelling framework explicitly accounts for the effects of crop rotations (on expected yields and variable input uses) on farmers' acreage choices.

The main objective of this paper is to propose an approach to estimate crop rotation effects,

which could in turn be used in the type of model of acreage choices proposed by Carpentier and Gohin (2015). Do to so, we consider two types of data: accountancy data and expert knowledge information.

As a matter of fact, estimating crop rotation effects on crop yields and input uses would be relatively easy with farm accountancy data such as the one we use if farmers crop sequence acreages were observed. Standard estimation approaches could be employed such as regressing farmers' choices observed at the crop level on crop sequence acreage choices describing how the acreage of a given crop is allocated to land areas with various cropping histories. Unfortunately, accountancy data do not describe farmers' crop sequence acreage choices. Thus, we need to devise an estimation approach to measure the effects of crop rotations without observing farmers' crop sequence choices, which is a challenging issue. Indeed, the lack of data, as regard to farmers' crop sequence choice in particular, likely is the primary factor explaining why crop rotation effects are rarely estimated. The estimation procedures considered in this paper are original: these are the first aimed to estimate crop rotation effects on farmers' yield and input use levels based on farm accountancy data only, *i.e.* without any quantitative information on farmers' crop sequence choices.

Our estimation approach aims at estimating crop rotation effects based on standard statistical criteria while simultaneously recovering the crop sequence choices that are consistent with the observed crop acreages and the estimated crop rotation effects.

From a practical viewpoint, our estimation procedure solve Bi-level Programming (BLP) problems in the parameters describing the effects of crop rotations on crop yields and input uses. The upper level problem consists of optimizing, in these crop rotation parameters, a statistical criterion based on the observed crop yield and input use levels and on the crop sequence acreages recovered at the lower level. The lower level recovers the crop sequence acreages assuming that they are optimally chosen by farmers given their observed crop

acreage choices and the estimated crop rotation parameters. Expert knowledge information is used to define constraints aimed at ‘guiding’ our estimation procedures to relevant crop rotation effect estimates.

The rest of the paper is organized as follows. The proposed approach to estimate the crop rotation effects to be used in this type of model is presented in the next part. In the third part, we discuss empirical estimation issues and present the estimation results we obtain with a French data sample. Finally we conclude.

1. Proposed approach to estimate crop rotation effects

1.1. Assumptions and models

Accountancy data generally available to estimate economic models of acreage choices provide information about crop production choices for each farm ($i = 1, \dots, N$) at each time period ($t = 0, 1, \dots, T_i$): crop acreages ($a_{k,it}$), obtained yield levels ($y_{k,it}^0$), available land area (A_{it}), variable input use levels per crop ($x_{jk,it}$) and crop prices paid to farmers ($p_{k,it}^0$), for $k \in \mathcal{K}$, the crop set available to farmers.

In our model, farmers are assumed to allocate their total arable land area A to crops k , *i.e.* to choose the crop acreages $a_{k,it}$ for $k \in \mathcal{K}$ with $\sum_{k \in \mathcal{K}} a_{k,it} = A_{it}$.

We further assume that farmers are forward looking and choose their crop sequence acreage by considering the effects of previous crops on the production process of the current crops. They are thus assumed to consider crop sequence acreages, *i.e.* in year t they consider the acreage of crop k produced on land with previous crop m , $s_{mk,it}$, for $(m, k) \in \mathcal{K} \times \mathcal{K}$.

Because farmers cannot use more or less acreage of a previous crop than the crop acreages defined by their last crop acreage choice, the crop sequence acreage choices are necessarily constrained by crop rotation constraints. These crop rotation constraints state that the demand

for land with a given previous crop equals the past acreage of this crop, *i.e.* $\mathring{a}_{k \hat{I} \mathcal{K}} s_{mk,it} = a_{m,it-1}$ for $m \in \mathcal{K}$.

At each time period, farmers thus seek to maximize their gross revenue which is defined as the sum of the crop sequence expected gross margins $p_{mk,it}$ weighted by the corresponding crop sequence acreages. The crop sequence expected gross margins depend on the effects of the previous crops on the expected yield and variable input use levels of the current crops. They measure the expected economic benefits of the considered crop pairs. These expected gross margins are defined as:

$$p_{mk,it} = \mathbf{y}_{mk,it}' \mathbf{p}_{k,it},$$

where $\mathbf{y}_{mk,it} = (y_{mk,it}^0, -\mathbf{x}_{mk,it})$ is the expected netput level vector, with $y_{mk,it}^0$ the expected yield level of the crop sequence (m,k) and $\mathbf{x}_{mk,it} = (x_{j,mk,it} : j \in \mathcal{J})$ the expected level of the corresponding variable input use vector. $\mathbf{p}_{k,it} = (p_{k,it}^j : j \in \mathcal{J})$ is the vector of expected netput price levels

The expected netput level vectors $\mathbf{y}_{mk,it}$ are unobserved and thus need to be estimated for every considered crop sequence (m,k) .

Let $y_{mk,it}^j$ denote the quantity of netput j used/produced per unit of land by farmer i in year t for the crop sequence (m,k) .

The observed crop netput levels are sums of the corresponding crop sequence netput levels weighted by the corresponding crop sequence acreage shares, *i.e.* we have:

$$y_{k,it}^j = \mathring{a}_{m \hat{I} \mathcal{K}} z_{mk,it} y_{mk,it}^j \text{ for } j \in \mathcal{J} \text{ and } k \in \mathcal{K}.$$

where $z_{mk,it}$ is the acreage share of crop sequence (m,k) in the acreage of crop k decided by

farmer i in year t :

$$z_{mk,it} = \begin{cases} \frac{s_{mk,it}}{a_{k,it}} & \text{if } a_{k,it} > 0 \\ 0 & \text{if } a_{k,it} = 0 \end{cases} \quad \text{for } m \in \mathcal{K} \text{ and } k \in \mathcal{K}$$

The considered estimation problem is challenging as we aim to identify statistical characteristics of the crop sequence netput levels $y_{mk,it}^j$ without observing the crop sequence acreage shares $z_{mk,it}$. This implies that modelling assumptions need to be imposed as regards to the specification of the crop rotation effects to be estimated, the crop sequence acreage shares and the relationships linking these terms. We consider three modelling assumption sets aimed at allowing the estimation of the crop rotation effects of interest while simultaneously recovering the crop sequence acreage shares $z_{mk,it}$.

Crop yield and input use models and specification of the crop rotation effects

Crop rotation effects are defined as differences with respect to a reference previous crop. Considering crop $r(k)$ as the benchmark previous crop for crop k , the crop rotation effect of previous crop m on the netput j level of crop k is given by:

$$d_{mk,it}^j = y_{mk,it}^j - y_{r(k),it}^j$$

for farmer i in year t .

Our first modelling assumption defines a statistical model of $y_{k,it}^j$ as a function of the crop sequence acreage share vector $\mathbf{z}_{k,it} = (z_{mk,it} : m \in \mathcal{K})$ and of a vector crop rotation effects to be estimated $\boldsymbol{\delta}_{k,0}^j = (d_{mk,0}^j : m \in \mathcal{K})$,¹ with:

¹ The ‘0’ indices added to the crop rotation parameters indicate the ‘true’ value of this parameter. We seek to

$$y_{k,it}^j = b_{k,t,0}^j + \mathbf{z}_{k,it}' \boldsymbol{\delta}_{k,0}^j + e_{k,it}^j \text{ with } \begin{cases} d_{r(k)k,0}^j = 0 \\ E[e_{k,it}^j] = 0 \end{cases} \text{ for } (k,j) \in \mathcal{K} \times \mathcal{J}.$$

Indeed, we mostly assume that (a) the crop rotation effects $d_{mk,0}^j$ are constant across the considered sample and (b) the error term $e_{k,it}^j$ impacting the level of $y_{k,it}^j$ doesn't depend on the crop sequence acreage share vector $\mathbf{z}_{k,it}$. These assumptions are admittedly restrictive but relaxing them would significantly increase the complexity of an estimation problem that is already challenging. The normalization constraints $d_{r(k)k,0}^j = 0$ needs to be imposed because the elements of the crop sequence acreage share vector $\mathbf{z}_{k,it}$ sum to 1 by construction. This implies that the considered crop rotation effects are defined with crop $r(k)$ as the benchmark previous crop. The $b_{k,t,0}^j$ terms are year specific intercepts aimed at capture economic and weather effects. Note that this implies the following models:

$$y_{mk,it}^j = b_{k,t,0}^j + d_{mk,0}^j + e_{k,it}^j$$

for the crop sequence netput quantities.

Our estimation approach further assumes that the variable vector $\mathbf{z}_{k,it}$ is exogenous with respect to the $e_{k,it}^j$ error term.

Estimating the crop rotation parameter vectors $\boldsymbol{\delta}_{k,0}^j$ would be very easy under these assumptions if the crop sequence acreage share vectors $\mathbf{z}_{k,it}$ were observed. In that case, it would basically suffice to regress $y_{k,it}^j$ on $\mathbf{z}_{k,it}$, *i.e.* to solve the following (constrained) Ordinary Least Squares (OLS) minimization problem:

estimate the true value, $\boldsymbol{\delta}_{k,0}^j$, of the $\boldsymbol{\delta}_k^j$ parameters .

$$\min_{\delta_k^j} \sum_{(i,t) \in S} u_{k,it}^j(\mathbf{z}_{k,it}; \delta_k^j) \quad s.t. \quad d_{r(k)k}^j = 0$$

where $S = \{1, \dots, N\} \times \{1, \dots, T\}$. The $u_{k,it}^j(\mathbf{z}_{k,it}; \delta_k^j)$ terms are residual terms implicitly accounting for the estimation of the $b_{k,t,0}^j$ parameters:

$$u_{k,it}^j(\mathbf{z}_{k,it}; \delta_k^j) = y_{k,it}^j - N^{-1} \sum_{i=1}^N y_{k,it}^j - \sum_{k=1}^K \mathbf{z}_{k,it}' \delta_k^j - N^{-1} \sum_{i=1}^N \mathbf{z}_{k,it}' \delta_k^j$$

Crop sequence acreage choice models

Our second modelling assumption is a standard economic rationality assumption. It specifically aims to circumvent the fact that the sequence acreage share vectors \mathbf{z}_{it} are unobserved. It allows devising an approach to recover the unobserved crop sequence acreage choices of a farmer based on the observed crop acreage choices $(\mathbf{a}_{i,t}, \mathbf{a}_{i,t-1})$ given estimates of the crop rotation effects.

This rationality assumption defines the crop acreage share vector \mathbf{z}_{it} as a solution to a simple LP problem stating that farmers decide their crop sequence acreages by maximizing their expected returns at the farm level².

This problem, denoted as ‘problem $\mathbf{LP}_{i,t}$ ’, is defined as follows:

² This problem can be shown to derive from a dynamic optimization program of a forward looking farmer maximizing an objective function $\sum_{mk} p_{mk} s_{mk} - C(\mathbf{a})$ under specific constraint on acreages (Carpentier and Goin, 2015).

$$\text{Problem } \mathbf{LP}_{i,t} : \max_{\hat{\mathbf{z}}_{i,t}} \sum_{k \in \mathcal{K}} \sum_{m \in \mathcal{K}} \hat{z}_{mk,it} a_{k,it} (\boldsymbol{\delta}'_{mk} \mathbf{p}_{k,it}) \quad s.t. \\ \mathbf{IMP}_{i,t}(\hat{\mathbf{z}}_{it}) = \mathbf{0}, \mathbf{SHA}_{i,t}(\hat{\mathbf{z}}_{i,t}) \leq \mathbf{0}, \mathbf{ROT}_{i,t}(\hat{\mathbf{z}}_{it}) = \mathbf{0}$$

where:

$$\boldsymbol{\delta}_{mk} = (d_{mk}^j : j \in \mathcal{J}) \text{ and } \boldsymbol{\delta} = (\boldsymbol{\delta}_{mk} : (m,k) \in \mathcal{K}' \times \mathcal{K})$$

We describe successively the objective function and the constraint set involved in problem

$\mathbf{LP}_{i,t}$. Let assume that farmer i decides to allocate his previous crop acreage $\mathbf{a}_{i,t-1}$ to the current crop acreage $\mathbf{a}_{i,t}$. The expected profit of this farmer is given by:

$$\sum_{k \in \mathcal{K}} \sum_{m \in \mathcal{K}} a_{k,it} \hat{z}_{mk,it} p_{mk,it}$$

where his crop sequence acreage share choices are given by the $\hat{z}_{mk,it}$ terms and the expected profit of a unit of land of crop k with previous crop m is given by

$$p_{mk,it} = \sum_{j \in \mathcal{J}} p_{k,t}^j b_{k,t,0}^j + \sum_{j \in \mathcal{J}} p_{k,t}^j d_{mk,0}^j = \sum_{j \in \mathcal{J}} p_{k,t}^j b_{k,t,0}^j + \boldsymbol{\delta}'_{mk} \mathbf{p}_{k,it}.$$

Using this definition of $p_{mk,it}$ allows to rewrite the expected profit of farmer i in year t as:

$$\sum_{k \in \mathcal{K}} \sum_{m \in \mathcal{K}} a_{k,it} \sum_{j \in \mathcal{J}} p_{k,t}^j b_{k,t,0}^j + \sum_{k \in \mathcal{K}} \sum_{m \in \mathcal{K}} a_{k,it} \hat{z}_{mk,it} (\boldsymbol{\delta}'_{mk} \mathbf{p}_{k,it})$$

given that $\sum_{m \in \mathcal{K}} \hat{z}_{mk,it} = 1$ by construction. As a result, if farmer i chooses $\hat{\mathbf{z}}_{it}$ optimally from an

economic viewpoint then he maximizes in $\hat{\mathbf{z}}_{it}$ the objective function of problem $\mathbf{LP}_{i,t}$,

$$\sum_{k \in \mathcal{K}} \sum_{m \in \mathcal{K}} a_{k,it} \hat{z}_{mk,it} (\boldsymbol{\delta}'_{mk} \mathbf{p}_{k,it}).$$

Of course, the choice of $\hat{\mathbf{z}}_{it}$ is constrained by the ‘unfeasibility’ constraints $\mathbf{IMP}_{i,t}(\hat{\mathbf{z}}_{it}) = \mathbf{0}$,

the ‘acreage share’ constraints $\mathbf{SHA}_{i,t}(\hat{\mathbf{z}}_{i,t})^3 \mathbf{0}$ and the ‘crop rotation’ constraints $\mathbf{ROT}_{i,t}(\hat{\mathbf{z}}_{i,t}) = \mathbf{0}$.

The solutions in $\hat{\mathbf{z}}_{i,t}$ to problem $\mathbf{LP}_{i,t}$ may be multiple, implying that the solution to this problem is defined as a solution set $\mathcal{Z}_{i,t}(\boldsymbol{\delta})$. Condition $\mathbf{z}_{i,t} \hat{=} \mathcal{Z}_{i,t}(\boldsymbol{\delta})$ simply states that the crop sequence acreage $\mathbf{z}_{i,t}$ is optimal for farmer i in year t given his current and previous crop acreages $(\mathbf{a}_{i,t}, \mathbf{a}_{i,t-1})$ and the crop rotation effects measured by $\boldsymbol{\delta}$.

Solving problem $\mathbf{LP}_{i,t}$ allows recovering estimates of $\mathbf{z}_{i,t}$ from the crop sequence acreages $(\mathbf{a}_{i,t}, \mathbf{a}_{i,t-1})$ based on the crop rotation effect vector $\boldsymbol{\delta}$. This recovery process is an essential component of the approach we propose for estimating $\boldsymbol{\delta}$. Yet, defining estimates of $\mathbf{z}_{i,t}$ as a solution to an optimization problem is appealing mostly from a theoretical viewpoint.

The Karush-Kuhn-Tucker (KKT) conditions characterizing the solution set $\mathcal{Z}_{i,t}(\boldsymbol{\delta})$ provide another, often practically more convenient, characterization of $\mathbf{z}_{i,t}$. They define $\mathbf{z}_{i,t}$ as a solution to a set of equality and inequality conditions. These KKT conditions are compactly denoted by the inequality condition:

$$\mathbf{KKT}_{i,t}(\hat{\mathbf{z}}_{i,t}; \boldsymbol{\delta}; \hat{\boldsymbol{\mu}}_{i,t}, \hat{\boldsymbol{\lambda}}_{i,t}) \preceq \mathbf{0}.$$

This condition is parameterized by $\hat{\boldsymbol{\mu}}_{i,t}$, the Lagrange multiplier (LM) vector related to the ‘crop rotation’ constraints $\mathbf{ROT}_{i,t}(\hat{\mathbf{z}}_{i,t}) = \mathbf{0}$, and $\hat{\boldsymbol{\lambda}}_{i,t}$, the LM vector related the other constraints, *i.e.* $\mathbf{IMP}_{i,t}(\hat{\mathbf{z}}_{i,t}) = \mathbf{0}$ and $\mathbf{SHA}_{i,t}(\hat{\mathbf{z}}_{i,t})^3 \mathbf{0}$. The following equivalence holds:

$$\begin{aligned} & \mathbf{z}_{i,t} \hat{\mathbf{1}} \mathbf{z}_{i,t}(\boldsymbol{\delta}) \\ & \mathbf{c} \\ & \mathbf{KKT}_{i,t}(\mathbf{z}_{i,t}; \boldsymbol{\delta}; \boldsymbol{\mu}_{i,t}, \boldsymbol{\lambda}_{i,t}) \leq \mathbf{0} \end{aligned}$$

where $(\boldsymbol{\mu}_{i,t}, \boldsymbol{\lambda}_{i,t})$ are optimal values of the $(\hat{\boldsymbol{\mu}}_{i,t}, \hat{\boldsymbol{\lambda}}_{i,t})$ LM vector.

The optimal Lagrange multiplier vector $\boldsymbol{\mu}_{i,t}$ is uniquely defined when the crop rotation constraints involved in problem $\mathbf{LP}_{i,t}$ are not redundant. Indeed, redundant crop rotation constraints can be eliminated *a priori*, based on characteristics of the crop acreages $(\mathbf{a}_{i,t}, \mathbf{a}_{i,t-1})$ and of the set of ‘unfeasible’ or ‘unwarranted’ crop sequences defined by the experts. These possible redundancies are assumed to be removed from the ‘crop rotation’ constraint $\mathbf{ROT}_{i,t}(\hat{\mathbf{z}}_{it}) = \mathbf{0}$.

Importantly, even in cases where $\mathbf{z}_{i,t}$ is uniquely defined it exhibits salient discontinuities in the parameter vector $\boldsymbol{\delta}$. These discontinuities underlie significant empirical issues arising when estimating $\boldsymbol{\delta}$.

Carpentier and Gohin (2014) showed that adding an entropic perturbation term to the objective function of problem $\mathbf{LP}_{i,t}$ allows alleviating these discontinuity issues. They showed that the solution in $\hat{\mathbf{z}}_{i,t}$ to the following perturbed version of $\mathbf{LP}_{i,t}$:

$$\begin{aligned} \text{Problem } \mathbf{SmLP}_{i,t}^r : \max_{\hat{\mathbf{z}}_{i,t}} & \sum_{k \in \mathcal{K}} \sum_{m \in \mathcal{M}} \hat{z}_{mk,it} a_{k,i,t} (\boldsymbol{\delta}'_{mk} \mathbf{p}_{k,it}) - \rho_{i,t}(\hat{\mathbf{z}}_{it}; r) \\ & s.t. \\ & \mathbf{IMP}_{i,t}(\hat{\mathbf{z}}_{it}) = \mathbf{0}, \mathbf{SHA}_{i,t}(\hat{\mathbf{z}}_{i,t})^3 \leq \mathbf{0}^3, \mathbf{ROT}_{i,t}(\hat{\mathbf{z}}_{it}) = \mathbf{0} \end{aligned}$$

where:

$$\rho_{i,t}(\hat{\mathbf{z}}_{it}; r) = r^{-1} \sum_{k \in \mathcal{K}} \sum_{m \in \mathcal{M}} \hat{a}_{k,i,t} \hat{z}_{mk,it} \ln \hat{z}_{mk,it}$$

is particularly convenient from an empirical viewpoint. This solution, denoted by $\mathbf{z}_{i,t}^r$, is unique. It also converges to a point in $\mathcal{Z}_{i,t}(\boldsymbol{\delta})$ as the (positive) perturbation parameter r grows to infinity, indicating that $\mathbf{z}_{i,t}^r$ can be considered as a reliable approximate solution to problem $\mathbf{LP}_{i,t}$ when r is sufficiently large.

Carpentier and Gohin (2014) also showed that $\mathbf{z}_{i,t}^r$ can be defined as a particularly well-behaved function of $(\boldsymbol{\delta}, \hat{\boldsymbol{\mu}}_{i,t}; r)$. This function, denoted by $\mathbf{z}_{i,t}^0(\boldsymbol{\delta}, \hat{\boldsymbol{\mu}}_{i,t}; r)$, is defined in analytical closed form and smooth – continuously differentiable at will – in $(\boldsymbol{\delta}, \hat{\boldsymbol{\mu}}_{i,t})$. This ‘smoothness’ property is particularly valuable when this function is used to construct objective functions of optimization problem to be solved in $(\boldsymbol{\delta}, \hat{\boldsymbol{\mu}}_{i,t})$ since it allows using standard, gradient-based, optimization algorithms.

The solution to problem $\mathbf{SmLP}_{i,t}^r$ is obtained from $\mathbf{z}_{i,t}^0(\boldsymbol{\delta}, \hat{\boldsymbol{\mu}}_{i,t}; r)$ when the crop rotation constraint Lagrange multiplier $\boldsymbol{\mu}$ is optimal, *i.e.* we have $\mathbf{z}_{i,t}^r = \mathbf{z}_{i,t}^0(\boldsymbol{\delta}, \boldsymbol{\mu}_{i,t}^r; r)$ when $\boldsymbol{\mu}_{i,t}^r$ is the optimal value of the Lagrange multiplier vector related to the ‘crop rotation’ constraints $\mathbf{ROT}_{i,t}(\hat{\boldsymbol{\mu}}_{i,t}) = \mathbf{0}$. Carpentier and Gohin (2014) also showed that $\boldsymbol{\mu}_{i,t}^r$ is the unique solution in $\hat{\boldsymbol{\mu}}_{i,t}$ to the equation system $\mathbf{ROT}_{i,t}(\mathbf{z}_{i,t}^0(\boldsymbol{\delta}, \hat{\boldsymbol{\mu}}_{i,t}; r)) = \mathbf{0}$.

Considering problem $\mathbf{SmLP}_{i,t}^r$ instead of problem $\mathbf{LP}_{i,t}$ is appealing for alleviating the discontinuity and multiple solution issues impairing the solutions to problem $\mathbf{LP}_{i,t}$.

Expert knowledge constraints and constraints based on the observed crop acreages

Expert knowledge is useful for recovering crop rotation effects on crop yield and variable input use levels when crop sequences are unobserved. Expert knowledge can be obtained in various forms, *e.g.*, quantitative or qualitative information on the crop rotation effects, or on farmers' crop sequence choices. Yet, experts are unlikely to deliver comprehensive quantitative information, whether on the crop rotation effects or on farmers' crop sequence choices. In most cases quantitative information is not available. For example, farmers' crop input uses are often poorly documented while crop sequence yields are only documented with experimental data.

We decided to gather qualitative information from experts – agronomists, agricultural scientists or extension agents – by directly questioning them. Experts were asked to rank the crop sequence average netput levels for a given crop, *i.e.* to rank the $y_{mk}^j = E\left[\sum_{it} y_{mk,it}^j\right]$ terms for $m \in \mathcal{M}$ for every netput $j \in \mathcal{J}$ and crop $k \in \mathcal{K}$. The obtained rankings allow defining the best, second best, *etc.*, previous crop(s) for any crop $k \in \mathcal{K}$ according to the criteria defined by the quantities of netput j for $j \in \mathcal{J}$. Experts were also asked to indicate whether the considered crop sequence are infeasible (due to, *e.g.*, incompatible crop biological cycles) or strongly unwarranted (due to, *e.g.*, dramatic pest or disease risks) in the considered geographical area.

Rankings were considered because netput quantity rankings were expected to be less demanding for experts than netput quantity estimates.

We aim to estimate the crop rotation parameters δ_0 of the models presented below based on estimates of farmers' unobserved sequence acreage shares \mathbf{z}_{it} , where $\mathbf{z}_{it} = (\mathbf{z}_{k,it} : m \in \mathcal{M})$ and $\delta_0 = (\delta_{k,0}^j : (k,j) \in \mathcal{K} \times \mathcal{J})$. The expert knowledge gathered from agricultural scientists and extension agents and the observed crop acreages $(\mathbf{a}_{i,t}, \mathbf{a}_{i,t-1})$ allow defining constraints to

be satisfied by the crop rotation parameters δ_0 or farmers' unobserved sequence acreage shares \mathbf{z}_{it} . We use these conditions to define constraints to be satisfied by any candidate value of the crop rotation parameters δ and of farmers' crop sequence acreage choice $\hat{\mathbf{z}}_{it}$.

Crop sequences considered as 'unwarranted' or 'unfeasible' by experts imply nullity restrictions on the corresponding elements of $\hat{\mathbf{z}}_{it}$ for any observation of the sample. Null crop acreages in the observed crop acreages $(\mathbf{a}_{i,t}, \mathbf{a}_{i,t-1})$ imply nullity restrictions on the corresponding elements of $\hat{\mathbf{z}}_{it}$. Crop sequence (m,k) is 'unfeasible' for farmer i in year t if $a_{k,it} = 0$ or $a_{m,it-1} = 0$. The nullity restrictions imposed on \mathbf{z} described above define the 'unfeasibility' constraints $\mathbf{IMP}_{i,t}(\hat{\mathbf{z}}_{it}) = \mathbf{0}$.

Crop rotation effects rankings define inequality restrictions on elements of $\delta_{k,0}^j$. Other constraints can also be defined, *e.g.* by imposing specific inequality restrictions on expected crop sequence gross margins. Such conditions impose ranking restrictions on elements of $\delta_{k,0}^j$ that can be described by inequality restrictions. These define the following 'expert knowledge' constraints $\mathbf{EXP}_{i,t}(\delta) \leq \mathbf{0}$. We assume here that these constraints also include the normalization constraints necessarily imposed on δ .

The crop rotation constraints imposed on $\hat{\mathbf{z}}_{it}$ by the crop acreages $(\mathbf{a}_{i,t}, \mathbf{a}_{i,t-1})$ are given by the conditions $\sum_{k \in \mathcal{K}} a_{k,it} \hat{z}_{mk,it} = a_{m,it-1}$. These conditions define the 'crop rotation' constraints $\mathbf{ROT}_{i,t}(\hat{\mathbf{z}}_{it}) = \mathbf{0}$.

Finally, the z_{mk} terms being defined as the share of crop sequence (m,k) in the crop k acreage, the elements of $\hat{\mathbf{z}}_{it}$ must satisfy technical conditions: the non-negativity constraints $\hat{z}_{mk,it} \geq 0$

and the crop acreage use constraint $\sum_{m \in \mathcal{M}} \hat{z}_{mk,it} = 1$. These conditions on \mathbf{z} define the ‘acreage share’ constraints $\mathbf{SHA}_{i,t}(\hat{\mathbf{z}}_{i,t})^3 \geq \mathbf{0}$.

2.3. General estimation approach

Our set up provides three main elements for devising an estimation approach of the crop rotation effects on the netput quantity levels:

- (a) *statistical models describing* how the observed crop *netput quantities* $y_{k,i,t}$ depend on farmers’ unobserved crop sequence crop share choices $\mathbf{z}_{k,i,t}$ and on the crop rotation parameters $\boldsymbol{\delta}_{k,0}^j$ to be estimated,
- (b) *procedures to obtain that crop sequence crop share choices* $\mathbf{z}_{i,t}$ consistent with the estimated values of the crop rotation parameters $\boldsymbol{\delta}_0$ and the observed crop acreage choices $(\mathbf{a}_{i,t}, \mathbf{a}_{i,t-1})$, based on standard or perturbed LP problems,
- (c) ‘*expert knowledge*’ *constraints* to be satisfied by any candidate value of the crop rotation parameters $\boldsymbol{\delta}_0$, and crop sequence acreage shares $\mathbf{z}_{i,t}$.

Our general estimation approach combine these components in the following theoretical bi-level programming (BLP) estimation problem:

$$\text{Problem U : } \min_{\boldsymbol{\delta}} \sum_{k \in \mathcal{K}} \sum_{j \in \mathcal{J}} \sum_{i \in \mathcal{I}} \sum_{t \in \mathcal{T}} u_{k,i,t}^j(\mathbf{z}_{k,i,t}; \boldsymbol{\delta}_k^j) \frac{y_{k,i,t}}{\bar{y}_{k,i,t}} \quad s.t. \quad \mathbf{EXP}_{i,t}(\boldsymbol{\delta}) \leq \mathbf{0}, \mathbf{z}_{i,t} \in \mathcal{Z}_{i,t}(\boldsymbol{\delta}), (i,t) \in \mathcal{I} \times \mathcal{T}$$

$$\text{Problem } \mathbf{L}_{i,t} : \quad Z_{i,t}(\boldsymbol{\delta}) = \arg \max_{\hat{\mathbf{z}}_{i,t}} \sum_{k \in \mathcal{K}} \sum_{m \in \mathcal{M}} \hat{z}_{mk,it} a_{k,i,t} (\boldsymbol{\delta}'_{mk} \mathbf{p}_{k,it}) \quad s.t. \\ \mathbf{IMP}_{i,t}(\hat{\mathbf{z}}_{it}) = \mathbf{0}, \mathbf{SHA}_{i,t}(\hat{\mathbf{z}}_{i,t})^3 \leq \mathbf{0}, \mathbf{ROT}_{i,t}(\hat{\mathbf{z}}_{it}) = \mathbf{0} \\ \text{for } (i,t) \in S.$$

Problem **U** is the upper level (master) problem. It seeks to minimize in the crop rotation parameter vector $\boldsymbol{\delta}$ the sum of squared residuals of the considered set of netput quantity models. This simple ‘pooled’ (across crops and netputs) OLS criterion yields a consistent estimator of $\boldsymbol{\delta}_0$ under the assumptions underlying our netput quantity models. More sophisticated estimators can be used, such as SUR-type estimators (due to the system structure of our netput model set). Nevertheless, the related efficiency gain is likely to be limited.

Problems $\mathbf{L}_{i,t}$ are the lower level (subordinate) problems. They aim to deliver ‘estimates’ of the crop sequence acreage vectors $\mathbf{z}_{i,t}$ that are used as explanatory variable vectors in the upper level (estimation) problem. According to our rationality assumption, these lower level problem are simple LP problems, *i.e.* problems $\mathbf{LP}_{i,t}$ for $(i,t) \in S$.

BLP problems cannot be solved directly, *i.e.* according to their theoretical formulation. We consider here a procedure based on transformations of the initial BLP problem

Mathematical programming with equilibrium constraints (MPEC) problem

BLP problems are generally solved by transforming it into a standard ‘one-level’ constrained optimization problem, according to the so-called ‘mathematical programming with equilibrium constraints’ (MPEC) approach. In this approach, the lower level problems $\mathbf{LP}_{i,t}$ are replaced by the optimality conditions characterizing their solutions. These KKT conditions are the ‘equilibrium constraints’ involved in the resulting MPEC problem. In our

case, adopting the MPEC approach consists of solving the following MPEC problem:

$$\begin{aligned} \min_{\delta, \hat{\mu}, \hat{\lambda}, \hat{z}} \quad & \sum_{k \in K} \sum_{j \in J} \sum_{(i,t) \in S} u_{k,i,t}^j (\mathbf{z}_{k,i,t}; \delta_k)^2 \frac{\bar{\theta}}{\theta} \\ \text{s.t.} \quad & \mathbf{EXP}_{i,t}(\delta) \leq \mathbf{0}, \mathbf{KKT}_{i,t}(\hat{\mathbf{z}}_{i,t}; \delta; \hat{\mu}_{i,t}, \hat{\lambda}_{i,t}) \leq \mathbf{0}, (i,t) \in S \end{aligned}$$

This problem is to be solved in the crop rotation parameter vector δ , crop sequence acreages

$\hat{\mathbf{z}} = (\hat{\mathbf{z}}_{i,t} : (i,t) \in S)$, ‘crop rotation’ constraint LM vectors $\hat{\mu} = (\hat{\mu}_{i,t} : (i,t) \in S)$ and other constraints LM vector $\hat{\lambda} = (\hat{\lambda}_{i,t} : (i,t) \in S)$.

The optimization problems to be solved in this approach are standard (‘one-level’) constrained optimization problems. Yet, these problems involve numerous discontinuous constraints (*i.e.*, the KKT conditions related to the non-negativity constraints of problems $\mathbf{LP}_{i,t}$ for $(i,t) \in S$) and are quite large.

Smooth mathematical programming with equilibrium constraints (SmMPEC) problem

The issues described above are alleviated when the lower level problems, problems $\mathbf{LP}_{i,t}$, are replaced by their perturbed versions, problems $\mathbf{SmLP}_{i,t}^r$, in the considered BLP problem. This is because the solutions in $\hat{\mathbf{z}}_{i,t}$ to problems $\mathbf{SmLP}_{i,t}^r$ are defined as known functions of $(\delta, \hat{\mu}_{i,t})$ by the $\mathbf{z}_{i,t}^o(\delta, \hat{\mu}_{i,t}; r)$ terms. Applying the MPEC approach, the BLP obtained by substituting problems $\mathbf{SmLP}_{i,t}^r$ for problems $\mathbf{LP}_{i,t}$ yields the following Smooth mathematical programming with equilibrium constraints (SmMPEC) problem:

$$\min_{\delta, \hat{\mu}} \sum_{k \in \mathcal{K}} \sum_{j \in \mathcal{J}} \sum_{(i,t) \in \mathcal{S}} u_{k,i,t}^j \left(\mathbf{z}_{i,t}^o(\delta, \hat{\mu}_{i,t}; r); \delta_k^j \right)^2 \frac{\bar{\sigma}}{\bar{\sigma}} \quad s.t. \\ \mathbf{EXP}_{i,t}(\delta) \leq \mathbf{0}, \mathbf{ROT}_{i,t}(\mathbf{z}_{i,t}^o(\delta, \hat{\mu}_{i,t}; r)) = \mathbf{0}, (i,t) \in \mathcal{S}$$

2. Empirical results

3.1. Data

Recovering the crop rotation parameters is an estimation problem that requires a data sample with sufficient size. As a result, we need to consider many observations. Then, farmers consider the entire crop set when deciding their crop sequence acreages. Hence, we need to consider a large crop set, the one containing the crops with the largest acreage shares in our sample. Furthermore, farmers are assumed to optimize their crop sequence acreage shares according to a profit maximization problem. This implies that we need to consider the main netputs, *i.e.* crop yields, fertilizer uses and pesticide uses. Finally, the considered estimation problems cannot be solved without ‘expert knowledge’ constraints on the crop rotation parameters and on the feasibility of crop sequences.

Accountancy data

This study is based on a dataset that is an unbalanced panel data sample with 487 observations describing the production choices of around 150 French arable crop producers located in the Marne *département*³ over the years 2011 to 2014. It was obtained from CDER, an accountancy firm, and provides detailed information on crop production for each farm: acreages, yields, crop prices at the farm gate and cost accounting (*i.e.*, variable input uses are reported at the crop level). The aggregated variable input price indices were constructed with the indices made available at the regional level by the French Department of Agriculture.

³ A *département* is a relatively small French territorial division (8162 km² for the Marne *département*).

Simple Paasche's indices were used for aggregating pesticides on the one hand, and fertilizers on the other hand. Pesticide uses and fertilizer uses are measured in €2010, and the corresponding price indices are constructed with 2010 as the base year.

The sampled farms mostly produce at least three crops selected from a seven crop set: winter wheat, spring barley, grain maize, rapeseed, sugar beet, alfalfa and protein pea.⁴

Table 1. Summary statistics, 2011 – 2014 averages

	Winter wheat	Spring barley	Grain maize	Rape-seed	Sugar beet	Alfalfa	Protein pea
Average yield (t/ha)	8.96	6.96	9.41	3.99	90.32	11.99	4.62
Average fertilizer use (€2010/ha)	252	187	197	240	307	267	81
Average pesticide use (€2010/ha)	186	109	112	215	280	69	157
Average gross margin (€/ha)	975	859	969	956	1830	663	857
Subsidies (€/ha)	0	0	0	0	0	185	125
Average acreage share when produced	0.357	0.188	0.148	0.170	0.156	0.108	0.062
Average acreage share	0.357	0.177	0.066	0.166	0.131	0.075	0.029
Production frequency (%)	100%	94%	45%	98%	84%	69%	46%
Previous average acreage share when produced	0.353	0.190	0.145	0.169	0.154	0.108	0.067
Previous average acreage share	0.353	0.179	0.066	0.164	0.129	0.077	0.033
Previous production frequency (%)	100%	94%	45%	97%	84%	71%	49%

The standard summary statistics displayed in Table 1 reveal that the sampled farmers' always produce wheat and almost always produce barley, rapeseed and, albeit to a lesser extent, sugar beet. Slightly less than half of the sampled farmers produce peas and maize. Around 69% produce alfalfa thanks to the outlet provided by local factories that dry alfalfa for feed.

The Marne *département* is among the most productive areas for arable crops in France, as shown by the average yield levels of wheat, nearly 9 t/ha, and of sugar beet, more than 90 t/ha, over the 2011-2014 period. Sugar beet is the most intensive crop in pesticide and fertilizer uses. Rapeseed is the second most intensive.

Sugar beet is the most profitable crop, with an average gross margin that is roughly twice that of the other crops, around 1830 €/ha. This is mostly due to the subsidized sugar beet quotas

⁴ These selection criteria are not as stringent as they may appear. During the considered period the acreages of the considered crops cover more than 90% of the arable land devoted to crops in the considered area.

that will be part of the CAP until 2017. Barley and alfalfa are the least profitable ones, with gross margins around 850 €/ha. Wheat, rapeseed and protein (including subsidies for this later crop) have comparable gross margin levels, around 950€/ha. These gross margin differences are important for crop sequence choices. If two crops compete for a given previous crop because it is well suited for these crops, then the land area with these previous crops is likely to be devoted to the crop with the largest gross margin.

The current and previous average acreage shares tend to show that the crop acreage shares do not vary much – or at least vary only slowly – across years.

Expert knowledge data

We have two sources of expert knowledge. The first one are the crop pair ratings found in the users' notice of the RPG Explorer software (Levavasseur *et al*, 2015) that allows exploiting the Integrated Administration and Control System (IACS) datasets. The ratings corresponding to the crop set considered in this study are reported in Table 2. These ratings lie from 0 to 10 in a given column, *i.e.* they rate previous crops for a given crop and 10 indicates the best crop. It is difficult to figure out how the different aspects of the production of a crop are accounted for – or weighted – in the given rates: yield levels, input use levels, gross margins, etc. Yet, these ratings can be interpreted as summaries of these different aspects. They were essentially used to define nullity constraints on the acreages of the crop sequence acreages that are unfeasible, unwarranted or characterized by low ratings. Additional rankings obtained from interviewed experts were used to define constraints on the crop sequence gross margins, *i.e.* for ordering the $\mathbf{p}'_{k,i,t} \boldsymbol{\delta}_{mk}$ terms for $m \in \mathcal{K}$. These experts are agricultural scientists specialized in arable crops. They supplied rankings of the previous crops of the seven considered crops with respect to three different criteria – expected yields, fertilizer uses and pesticide uses. These rankings are reported in Tables 7a–c.

Table 2. Previous crop ratings of the RPG Explorer notice

		Current crops						
		Winter wheat	Spring barley	Grain maize	Rape-seed	Sugar beet	Alfalfa	Protein pea
Previous crops	Winter wheat	4	8	10	8	10	10	10
	Spring barley	6	4	10	10	10	10	8
	Grain maize	8	8	6	2	6	6	6
	Rapeseed	10	10	4	1	4	10	4
	Sugar beet	10	10	4	2	2	5	6
	Alfalfa	10	1	10	10	6	10	4
	Protein pea	10	6	6	4	6	10	1

Interpretation: 10 = excellent ; ... ; 0 = very bad

Table 3a. Interviewed experts' previous crop rankings - Criterion: expected yield

		Previous crops		Current crops				
		Winter wheat	Spring barley	Grain maize	Rape-seed	Sugar beet	Alfalfa	Protein pea
Previous crops	Winter wheat	5	2	1	1	1	1	1
	Spring barley	4	5	1	1	1	1	1
	Grain maize	4	1	3	0	3	1	1
	Rapeseed	2	1	1	0	2	2	2
	Sugar beet	3	3	2	0	0	3	3
	Alfalfa	1	4	1	2	4	(1)	0
	Protein pea	1	4	1	2	4	0	0

Interpretation: 0 = Unfeasible or strongly unwarranted ; 1 = highest ; 2 = second highest ; etc

Table 3b. Interviewed experts' previous crop rankings - Criterion: expected fertilizer use

		Previous crop		Current crops				
		Winter wheat	Spring barley	Grain maize	Rape-seed	Sugar beet	Alfalfa	Protein pea
Previous crops	Winter wheat	4	4	4	2	2	2	1
	Spring barley	4	4	4	2	2	2	1
	Grain maize	4	4	4	0	2	2	1
	Rapeseed	3	3	3	0	1	2	1
	Sugar beet	2	2	2	0	0	2	1

Alfalfa	1	1	1	1	1	1	0
Protein pea	1	1	1	1	1	0	0

Interpretation: 0 = Unfeasible or strongly unwarranted ; 1 = lowest ; 2 = second lowest ; etc

Table 3c. Interviewed experts' previous crop rankings - Criterion: expected pesticide use

		Current crops						
		Winter wheat	Spring barley	Grain maize	Rape-seed	Sugar beet	Alfalfa	Protein pea
Previous crops	Winter wheat	3	2	2	1	1	2	1
	Spring barley	2	3	2	1	1	2	1
	Grain maize	2	2	3	1	1	2	1
	Rapeseed	1	1	1	0	1	2	1
	Sugar beet	1	1	1	0	1	2	1
	Alfalfa	1	1	1	0	0	(1)	0
	Protein pea	1	1	1	1	1	0	0

Interpretation: 0 = Unfeasible or strongly unwarranted ; 1 = lowest ; 2 = second lowest ; etc

The main advantage of these per criterion rankings lies in the fact that they allow for designing constraints on the crop rotation parameters to be estimated, *i.e.* on the elements of the d_{mk}^j terms for $m \in \mathcal{K}$. Moreover, the expert rankings are established according to well defined criteria whereas the RGP Explorer ratings are established on criteria that are unclear. We are thus more confident in the constraints obtained from the consulted experts than in those reported in the RPG Explorer notice.

The rankings obtained from the interviewed experts are generally consistent with those contained in the RGP Explorer notice. Yet, notable differences appear with the rankings of the previous crops for barley and, to a lesser extent, for maize. According to the interviewed experts, legumes – alfalfa and protein pea – are among the best previous crops for barley whereas they are among the worst according to the RGP Explorer notice. Also, the interviewed experts assert the (pea, alfalfa) crop sequence is unwarranted due to disease problems, notwithstanding that this sequence involves two legumes.

3.2. Estimation strategy

Expert knowledge constraints. The expert constraints are of two types. The ‘unfeasibility constraints’ impose nullity constraints on the acreage shares that are unfeasible or strongly unwarranted according the collected expert knowledge. The estimates, as well as their reliability, of crop rotation effects strongly depend on the considered set of unfeasibility constraints. First, we impose simple unfeasibility constraints. All monoculture crop sequences, *i.e.* the (k,k) crop sequences, are considered as unfeasible with corresponding null acreage shares (and, of course, unidentified corresponding crop rotation parameters). The sole exception is the (wheat,wheat) crop sequence that cannot be avoided in more than 10% of the sample observations. We also excluded the (alfalfa,pea) and (pea,alfalfa) sequences.

The ‘ranking constraints’ impose rankings on the crop rotation parameters, they can be imposed directly on the considered parameters or through ranking constraints imposed on the crop sequence gross margins. The two options have tested. Unfortunately, the estimates based on the constraints imposed on the crop rotation parameters are less precise than those based on the constraints imposed on the crop sequence margins. In particular, the constraints imposed on the crop rotation parameters yield crop rotation effect estimates that are either null or rather large in absolute value. We thus used information reported in Table 3a-3c to define the ranking constraints on crop sequence gross margins and imposed these constraints in our estimations.

We also consider a specific constraint on the (alfalfa,alfalfa) acreage shares since alfalfa is grown for two to three years in the Marne *département*. This constraint states that the sample mean of the share of alfalfa using alfalfa as previous crop must exceed, at least, 0.5. This constraint is defined at the sample level because the observation level constraints are likely to be unduly restrictive for many observations.

Solution approaches to bi-level programming problems. The smooth MPEC formulation of the considered crop rotation effect estimation problems is used to solve the related bi-level programming problems. Indeed, as shown in the previous section, this approach yields a formulation of the bi-level programming problem that is easy to interpret and well behaved. It relies on a basic OLS estimation criterion. It replaces the LP problems aimed at obtaining the optimal acreage share vectors $\mathbf{z}_{i,t}$ by their counterparts with entropic perturbations. As a result, these optimal acreage share vectors are obtained analytical closed form solutions $\mathbf{z}_{i,t}^r(\boldsymbol{\delta}_k, \boldsymbol{\mu}_{i,t}; r)$ up to the optimal level of the crop rotation Lagrange multiplier vector $\boldsymbol{\mu}_{i,t}$ that is obtained by solving the problem along with the interest parameter vector $(\lambda, \boldsymbol{\delta})$.

Formally, the smooth MPEC formulation is given by:

$$\begin{aligned}
& \min_{\delta, \lambda, \mu} \sum_{k \in \mathcal{K}} \sum_{j \in \mathcal{J}} \sum_{(i,t) \in \mathcal{IS}} \hat{e}_{k,i,t}^{OLS} (\lambda_k^j, \delta_k^j; \mathbf{z}_{i,t}^r (\delta_k, \mu_{i,t}; r))^2 \frac{\bar{\sigma}_{k,i,t}^2}{\bar{\sigma}} \\
& \quad s.t. \\
& \quad \mathbf{z}_{mk,i,t}^r (\delta_k, \mu_{i,t}; r) = \frac{\exp \left(\mathbf{p}_{k,i,t}' (\delta_{mk} - m_{m,i,t}) \right)}{\sum_{n \in \mathcal{K}} \exp \left(\mathbf{p}_{k,i,t}' (\delta_{nk} - m_{n,i,t}) \right)}, \text{ for } (m,k,i,t) \in \mathcal{K}^2 \times \mathcal{S} \\
& \quad \text{(Optimal crop sequence acreage choices, with entropic perturbation)} \\
& \quad \mathbf{a}_{i,t}' \mathbf{z}_{m,i,t}^r (\delta_{k,0}, \mu_{i,t}; r) = a_{m,i,t-1}, \text{ for } m \in \mathcal{K} \\
& \quad \text{(Crop rotation constraints)} \\
& \quad d_{r(k)k}^j = 0 \text{ and } \mathbf{1}_T' \lambda_k^j = 0, \text{ for } (k,j) \in \mathcal{K} \times \mathcal{J} \\
& \quad m_{r(k)k}^j = 0, \text{ for } (k,i,t) \in \mathcal{K} \times \mathcal{S} \\
& \quad \text{(Normalization constraints)} \\
& \quad \mathbf{z}_{mk,i,t}^r (\delta_k, \mu_{i,t}; r) = d_{mk}^j = 0 \text{ if } (m,k) \text{ is unfeasible, for } (j,i,t) \in \mathcal{J} \times \mathcal{S} \\
& \quad \text{(Crop sequence unfeasibility constraints)} \\
& \quad \mathbf{p}_{i,t}' \delta \leq 0, \text{ for } (i,t) \in \mathcal{S} \\
& \quad \text{(Crop sequence gross margin rankings)} \\
& \quad \sum_{(i,t) \in \mathcal{IS}} d_{n,i,t} d_{n,i,t-1} \frac{\bar{\sigma}_{n,i,t}^2}{\bar{\sigma}} \sum_{(i,t) \in \mathcal{IS}} \mathbf{z}_{mk,i,t}^r (\delta_k, \mu_{i,t}; r)^3 \leq 0.5, \text{ for } n = \text{alfalfa} \\
& \quad \text{(Constraints on the sample mean of the (alfalfa, alfalfa) acreage shares)}
\end{aligned}$$

where $d_{k,i,t}$ is the indicator variable of the production of crop k by farmer i in year t :

$$d_{k,i,t} = \begin{cases} 1 & \text{if } a_{k,i,t} > 0 \\ 0 & \text{if } a_{k,i,t} = 0 \end{cases}$$

Standard (gradient based) algorithms perform well for solving the implied smooth nonlinear optimization problem, at least as long as the level of the perturbation parameter r is small. However, large levels of the perturbation parameter r are required for the solutions to the perturbed LP problems $\mathbf{z}_{i,t}^r (\delta_k, \mu_{i,t}; r)$ to be close to the solutions to the original LP problems. Because such large levels of r also generate overflow and underflow problems, the perturbed

problem was solved using a simple iterative process. It was first solved starting with a small level of r and then solved with increasing values of r until numerical convergence of the $\mathbf{z}_{i,t}^r(\boldsymbol{\delta}_k, \boldsymbol{\mu}_{i,t}; r)$ terms. At low levels of r the $\mathbf{z}_{i,t}^r(\boldsymbol{\delta}_k, \boldsymbol{\mu}_{i,t}; r)$ terms are close to the previous crop acreages $a_{m,i,t-1}$ for $(m,k) \in \mathcal{K}^2$ while they converge to an optimal solution to the LP problem that is assumed by farmer i in year t .

The NLP solver available from GAMS quickly solves the smooth MPEC problem by the search process of the optimal level of the perturbation parameter r significantly increases the computing time.

3.3. Estimation results

The obtained estimates of the crop rotation parameters are reported in Tables 8a–c. These estimates are relatively large for the yield crop rotation effects, in absolute value and with respect to the average crop yield levels. Their rankings were expected for wheat although the levels of the rotation effects related to the wheat yields are relatively large. The rankings of the crop rotation effects related to the barley yields are surprising. According to these rankings, the best previous crop of barley is wheat and using rapeseed or protein pea as previous crops induce losses around 1 t/ha.

The estimated crop rotation effects on fertilizer uses are surprising. In particular, according to these results using alfalfa or protein as previous crops induce significant increases in fertilizer uses, *e.g.* around 65 €/2005/ha for wheat, whereas these crops are legumes and capture significant amounts of atmospheric nitrogen that are available for the following crops. Indeed, these estimated effects on fertilizer uses seem to compensate the large estimated effects on the yield levels.

Table 4a. Estimated crop rotation effects on yield levels (t/ha),

		Current crops						
		Winter wheat	Spring barley	Grain maize	Rape-seed	Sugar beet	Alfalfa	Protein pea
Previous crops	Winter wheat	-1.51	0.00	-0.41	-0.01	-4.4	-0.21	0.00
	Spring barley	-1.51		-0.51	0.00	0.00	-1.31	-0.28
	Grain maize	-1.51	-1.38		-0.66	-7.5	-1.10	0.00
	Rapeseed	-0.49	-1.15	-1.92		-9.9	-2.36	-0.28
	Sugar beet	0.00	-0.08	0.00	-0.54		-0.94	-0.28
	Alfalfa	-0.35	-0.94	-0.75	0.02	-4.0	0.00	
	Protein pea	-0.19	-0.89	-0.34	-0.61	-5.5		
Average crop yield (t/ha)		8.96	6.96	9.41	3.99	93.2	11.99	4.62

Table 4b. Estimated crop rotation effects on fertilizer use levels (€/ha)

		Current crops						
		Winter wheat	Spring barley	Grain maize	Rape-seed	Sugar beet	Alfalfa	Protein pea
Previous crops	Winter wheat	0	4	14	39	58	36	5
	Spring barley	18		21	46	66	12	28
	Grain maize	32	0		0	0	22	0
	Rapeseed	60	16	9		65	0	28
	Sugar beet	64	47	80	57		109	28
	Alfalfa	94	33	0	96	41	46	
	Protein pea	69	37	35	36	111		
Average crop fertilizer use (€2005/ha)		252	187	197	240	307	267	81

Table 4c. Estimated crop rotation effects on pesticide use levels (€/ha)

		Current crops						
		Winter wheat	Spring barley	Grain maize	Rape-seed	Sugar beet	Alfalfa	Protein pea
Previous crops	Winter wheat	56	8	44	23	4	0	0
	Spring barley	33		26	18	0	13	17
	Grain maize	15	14		39	19	18	7
	Rapeseed	41	0	0		9	20	17
	Sugar beet	26	18	83	17		57	17
	Alfalfa	0	30	30	0	40	5	
	Protein pea	64	32	43	11	14		
Average crop pesticide use (€2005/ha)		186	109	112	215	280	69	157

The estimated crop rotation effects on pesticides uses were expected for wheat but they are surprising for barley. According to these estimates, to produce wheat with cereals as previous

crops increase pesticide uses while to produce barley with wheat as the previous is the best option with respect to pesticide uses.

These estimated crop rotation parameters directly impact the crop sequence gross margins. The sample means of the estimated differences in the crop sequence gross margins for the considered crops are reported in Table 5. These gross margin differences are significant in absolute value as well as with respect to the crop gross margin levels. The differences in the crop gross margins between the best and worst previous crops for the considered can amount to up to 20% of the crop gross margin of wheat or barley, and up to around 33% for alfalfa. Indeed, such large differences were unexpected.

Also, if the obtained rankings of the differences in crop sequence gross margins were those expected for wheat, they are surprising for barley and, to a lesser extent, for sugar beet. Wheat is considered as an excellent previous crop for sugar beet.

Table 5. Average estimated crop rotation effects on the crop gross margins (€/ha)

		Current crops						
		Winter wheat	Spring barley	Grain maize	Rape-seed	Sugar beet	Alfalfa	Protein pea
Previous crops	Winter wheat	-215	0	0	-1	-116	-2	0
	Spring barley	-214		-4	0	0	-92	-113
	Grain maize	-213	-248		-218	-144	-87	0
	Rapeseed	-95	-214	0		-283	0	0
	Sugar beet	0	-76	-63	-223		-220	-113
	Alfalfa	-72	-226	-16	-34	-121	0	
	Protein pea	-77	-225	-15	-220	-221		
Average crop gross margin (€/ha)		975	859	969	956	1830	663	867

Table 6 reports the sample means of the estimated crop sequence acreage shares, $z_{mk,i,t}$, when the considered crop sequences are feasible for the considered observation. *I.e.*, the (m,k) crop sequence is feasible for observation (i,t) if farmer i produced crop k in year t and crop m in year $t-1$. These sample means allow for identifying the preferred previous crops of the considered crops.

Table 6. Average estimated crop sequence acreage shares per crop, when feasible

		Current crops							Mean crop acreage share
		Winter wheat	Spring barley	Grain maize	Rape- seed	Sugar beet	Alfalfa	Pea	
Previous crops	Winter wheat	0.04	0.86	0.53	0.60	0.08	0.29	0.31	0.36
	Spring barley	0.01		0.23	0.33	0.81	0.06	0.03	0.18
	Grain maize	0.19	0.04		0.06	0.22	0.32	0.76	0.07
	Rape-seed	0.43	0.01	0.00		0.01	0.00	0.00	0.17
	Sugar beet	0.38	0.12	0.03	0.00		0.00	0.00	0.13
	Alfalfa	0.07	0.00	0.16	0.10	0.04	0.50		0.07
	Protein pea	0.15	0.00	0.28	0.00	0.02			0.03
	Mean crop acreage share	0.35	0.18	0.07	0.16	0.13	0.08	0.03	

The obtained results were generally expected. For instance, sugar beet and rapeseed are the preferred previous crops for wheat. More generally, the preferred previous crops of the cereal crops – wheat, barley and maize – are the rotation heads – rapeseed, pea, sugar beet and alfalfa – while the opposite also holds. Wheat appear to be an important previous crops for the other crops owing to its large acreage shares. Note also that the constraint on the sample mean of the (alfalfa,alfalfa) crop sequence acreage share binds at the obtained solution. Indeed, many ‘expert knowledge’ constraints also bind at the obtained solution. Indeed, the ‘expert knowledge’ constraints always play a crucial role in our results.

Table 7. Average estimated crop sequence acreage shares ($s_{mk,i,t} = z_{mk,i,t} a_{k,i,t}$)

		Current crops							Mean crop acreage share
		Winter wheat	Spring barley	Grain maize	Rape- seed	Sugar beet	Alfalfa	Pea	
Previous crops	Winter wheat	0.02	0.15	0.04	0.10	0.01	0.02	0.01	0.36
	Spring barley	0.01		0.01	0.05	0.10	0.00	0.00	0.18
	Grain maize	0.04	0.00		0.00	0.01	0.01	0.01	0.07
	Rapeseed	0.14	0.00	0.00		0.00			0.17
	Sugar beet	0.11	0.02	0.00	0.00		0.00	0.00	0.13
	Alfalfa	0.02	0.00	0.00	0.01	0.01	0.04		0.07
	Protein pea	0.03	0.00	0.01	0.00	0.00			0.03
	Mean crop acreage share	0.35	0.18	0.07	0.16	0.13	0.08	0.03	

Table 7 reports the sample means of the estimated crop sequence acreages as shares of the total farm arable land area, $s_{mk,i,t} = z_{mk,i,t} a_{k,i,t}$. These shares are automatically null when the considered crop sequences are unfeasible for the considered observation, *i.e.* we necessarily have $s_{mk,i,t} = 0$ if farmer i did not produce crop k in year t and/or did not produce crop m in year $t - 1$. Of course, we also have $s_{mk,i,t} = 0$ if farmer i considered in year t that better crop sequences than (m,k) could be used. These acreages allow for assessing the importance of crop sequence choices in the considered sample.

These acreage shares show that the crop sequences involving the major crops – *i.e.*, especially wheat and, to a lesser extent, rapeseed, barley and sugar beet in the considered sample – mechanically account for most land uses. As expected wheat is mostly produced following rapeseed (14.4%) and sugar beet (10.6%), barley after wheat (14.7%), rapeseed after wheat (9.8%) or barley (5.2%), sugar beet after barley (11.1%). Similarly, pea is mostly followed by wheat (2.5%). The use of alfalfa as a previous crop is more variable.

Many sample mean crop sequence acreage shares reported in Table 11 are very small, slightly or largely below 1%. This is good news for the entropic perturbation approach. While the $\mathbf{z}_{i,t}^r(\boldsymbol{\delta}_k, \boldsymbol{\mu}_{i,t}; r)$ terms lies in the (0,1) interval by construction, sufficiently large levels of r allow for obtaining acreage shares close to 0 (as well as close to 1 at the individual level). Similarly, in cases where the perturbation parameter is “too” small the crop sequence acreage shares $a_{k,i,t} \mathbf{z}_{i,t}^r(\boldsymbol{\delta}_k, \boldsymbol{\mu}_{i,t}; r)$ terms are close to the previous crop acreages $a_{m,i,t-1}$ for $m \in \mathcal{K}$. The estimated crop sequence acreage shares reported in Table 11 clearly show that this does not occur in our results.

Conclusion

In this paper we consider a BLP problem designed to estimate crop rotation effects on yield and input uses while simultaneously reconstructing farmers' unobserved crop sequences from farmers' observed (current and previous) crop acreages. Our estimation approach is based on a well-defined statistical background and relies on simple crop sequence yield and input use models and on an assumption stating that farmers are economically rational when deciding their crop sequence acreages.

BLP problems are complex problems and, because they are designed for solving involved estimation problems, the ones considered here are rather large. Moreover, specific features of crop sequence acreage choices imply specific estimation and, consequently, programming issues. Our estimation results suggest that the Smooth MPEC estimation procedure is promising for solving the considered BLP estimation problem. The 'smoothness' of the related optimization problem facilitates the estimation process monitoring.

References

- Carpentier, A., Gohin, A. (2015). On the economic theory of crop rotations: value of the crop rotation effects and implications on acreage choice modeling. Working Paper SMART – LERECO N°15-04.
- Carpentier, A. and A. Gohin (2014). Dynamic acreage choices. Selected paper, *14th Triennial Congress of the European Association of Agricultural Economists*. Ljubjana, Slovenia, August 2014.
- Eckstein Z. (1984). A Rational Expectations Model of Agricultural Supply. *Journal of Political Economy*, 92(1):1-19.
- Levavasseur F., Martin P. et Scheurer O. (2015). *RPG Explorer Version 1.8.37- Octobre*

2015. *Notice d'utilisation*. AgroParisTech, INRA et Institut Polytechnique Lasalle. 113p.
- Meynard, J.-M., A. Messéan, A. Charlier, F. Charrier, M. Farès, M. Le Bail, M.B. Magrini, 2013. *Freins et leviers à la diversification des cultures. Etude au niveau des exploitations agricoles et des filières*. Rapport d'étude, INRA, 226 p.
- Ozarem P.F., Miranowski J.A. (1994). A Dynamic Model of Acreage Allocation with General and Crop-Specific Soil Capital. *American Journal of Agricultural Economics*, 76(2):385-395.
- Tegene A., Huffman W.E., Miranowski J.A. (1988). Dynamic Corn Supply Functions: A Model with Explicit Optimization. *American Journal of Agricultural Economics*, 73(1):103-111.
- Thomas A. (2003). A Dynamic Model of On-Farm Integrated Nitrogen Management. *European Review of Agricultural Economics*, 30:439-460.
- Vitale, J.D., H. Djourra and A. Sidibé. (2009). Estimating the supply response of cotton and cereal crops in smallholder production systems: recent evidence from Mali. *Agricultural Economics*, 40:519–533.