On the economic value of the agronomic effects of crop diversification for farmers: Estimation based on farm cost accounting data

Ibirenoye Romaric Sodjahin[1], Fabienne Femenia[1], Obafèmi Philippe Koutchadé[1], Alain Carpentier[1]

[1]UMR 1302 SMART, INRAE, France

fabienne.femenia@inrae.fr

Selected Paper prepared for presentation at the 2022 Agricultural & Applied Economics Association Annual Meeting, Anaheim, CA; July 31-August 2

On the economic value of the agronomic effects of crop diversification for farmers: Estimation based on farm cost accounting data

Abstract

We estimate the effects of previous crops and crop acreage diversity on yields and chemical input uses. Our estimation approach relies on models of crop yields and input uses defined as systems of simultaneous equations featuring farm specific random parameters. We find significant and consistent effects of previous crops on yield levels but not on input uses, suggesting that farmers tend to downplay these effects when deciding their uses of chemical inputs. Our results also show that crop acreage diversity positively impacts yield levels and tends to induce reductions in pesticide uses. These effects are however limited from an economic viewpoint.

1. Introduction

The Common Agricultural Policy (CAP) recently put forward crop diversification at the farm level as a primary objective. The 2013 reform actually introduced a set of "crop diversification" obligations as eligibility criterion for farmers to receive the green direct payments[1] and the future CAP will include similar standards on farm crop acreages as part of its cross-compliance greening scheme. Current agri-environmental and climate schemes and future eco-schemes also aim to foster crop diversification in EU (European Union) farms (*e.g.*, Guyomard *et al*, 2020).

According to the EC (European Commission), the principle behind the greening obligations and congruent direct payments is "to remunerate farmers for their efforts to protect the environment and biodiversity" (EC, 2017). Yet, crop diversity can also yield on-farm benefits (*e.g.*, Ikerd, 1993; Lin 2011; Duru *et al*, 2015). As far as agronomic effects are concerned, these benefits can arise from three main channels: pre crop effects, crop rotation effects and spatial crop diversity effects. Previous crops are in fact expected to deliver pre crop effects to the crops that immediately follow them on the considered plot. Nitrogen surpluses delivered by legumes and break crop effects on biotic pressures are examples of these short run effects (*e.g.* Mortensen and Smith, 2020). In the medium run, crop rotation diversification is expected to increase the delivery of ecosystem services supporting agricultural production at the plot and farm scales. These ecosystem services notably include vegetation and natural enemy effects on pest and weed population regulations, pollination effects and improved soil structure, fertility and health (*e.g.*, Davis *et al*, 2012). If diversified crop rotations concern crop diversity across time on a given plot, they generally also imply crop acreage diversity across space at the farm level. Diversified crop acreages yield economic benefits that are well known in the economics literature, including those due to risk spreading, but crop diversity can also yield agro-

[1] These obligations were in fact not really constraining for most European farms (Louhichi *et al*, 2018).

ecological benefits whether it is implemented at the plot, at the farm or at the landscape level[2]. Disentangling crop rotation effects and spatial crop diversity effects is difficult with farm data since farmers simultaneously optimize crop rotation and crop acreage diversities. In what follows, we thus refer to both crop rotation and spatial crop diversity effects as "crop acreage diversity" effects.

In this article, we shed light on a pair of apparently contradictory facts: European farmers tend to stick to specialized crop acreages despite agronomic experiments tending to show that crop diversification could reduce chemical input uses while maintaining or even enhancing arable crop yield levels (*e.g.*, Duru *et al*, 2015; Thérond *et al*, 2017). Assessing the effects of both crop acreage diversity and pre crops on yield and input use levels based on farm data is of special interest in this context. In fact, since most studies considering crop acreage diversity or pre crop effects are based on agronomic experiments and focus on specific crop sequences or crop mixes[3], little is known on how crop diversification actually performs in commercial farms, and on how farmers perceive and make use of pre crop and crop acreage diversity effects.

Farmers may be reluctant to diversify their crop acreages and rotations for different reasons. They may expect diversified cropping mixes to be less profitable than specialized ones because typical diversification crops are less profitable than major crops, because they tend to understate the value of the positive agro-ecological effects induced by crop diversification (*e.g.*, Magrini *et al,* 2016; Watson *et al,* 2017), or because the need to monitor their crops to fully benefit from

---

[2] See, *e.g.*, Duru *et al* (2015) and Delaune *et al* (2021).

[3] Cropping system effects are mostly investigated through long run experiments in corn and soybean based systems in the US (*e.g.*, Davis *et al*, 2012; Hunt *et al*, 2017, Bowles *et al*, 2020, Feng *et al,* 2021) and in wheat based systems with a specific focus on canola grain legumes in Canada (*e.g.*, Zentner *et al*, 2002; Thiessen Martens *et al,* 2017; Smith *et al*, 2017; Liu *et al*, 2019; Khakbazan *et al* 2020). Experiments conducted in Europe mostly concern pre crop effects, with a particular focus on the effects of grain legumes on cereal yield and nitrogen fertilization levels (*e.g.*, Meynard *et al*, 2013; Bues *et al*, 2013; Watson *et al*, 2017; Hufnagel *et al*, 2020). Agronomic experiments considering the effects of crop diversification on pesticide use levels are relatively rare and generally focus on herbicide use levels (*e.g.*, Sharma *et al*, 2021).

diversification may induce implicit costs for them. The agri-food value chain organization can also hinder the adoption of diversified crop acreages by farmers when, for instance, suitable extension services are lacking or local outlets for diversification crops are insufficient. Such so-called socio-technical lock-ins are often put forward for explaining the low adoption of diversified crop acreages (*e.g.*, Magrini *et al,* 2016; Mortensen and Smith, 2020; Morel *et al,* 2020). Yet, policy instruments aimed to address lock-in issues substantially differ from instruments aiming to overcome lack of profitability, whether real or perceived. Distinguishing profitability and lock-in issues is thus crucial from an agri-environmental policy perspective and assessing the pre crop and crop acreage diversity effects on farm crop production and on farmer choices is an important step toward this overall objective.

The main purpose of our study is to assess pre crop and crop acreage diversity effects based on commercial farm data for, in turn, investigating whether farmers relying on diversified crop acreages tend to use less chemical inputs and whether farmers adjust their chemical inputs uses to the pre crop of the considered crops.

Most economic studies considering the effects of crop acreage diversity at the farm level focus on the productivity effects. They measure the effects of crop acreage diversity indicators on crop production value aggregates (*e.g.*, Omer *et al*, 2007; Di Falco and Chavas, 2008; Di Falco *et al*, 2010; Groom and Pereira Fontes, 2021), on expected return measures or/and on return risk measures (*e.g.*, Di Falco and Perrings, 2005; Bozzola and Smale, 2020).

Chavas and Kim (2007 and 2010), Chavas (2009) and Chavas and Di Falco (2012) propose to analyze the interest of crop diversification by relying on producer theory in various multi-output settings. Their frameworks they provide theoretically grounded measures of the benefits of crop diversification as well as useful decompositions of these measures, which highlight the effects of production complementarities across crops and are closely related to the agronomic effects considered in this article. These theoretical frameworks are however essentially static.

The recent study of Bareille and Letort (2018) measures the effects of crop acreage diversity on crop yield and chemical input use levels based on farm cost accounting data. They generally demonstrate positive effects of crop diversity on crop yield levels, and negative effects on chemical input uses and on yield variability. However, as most empirical analyses of crop diversity and probably due a lack of data on crop sequence acreages, pre crop effects are ignored in their study.

The approach we propose here enables to disentangle pre crop and crop acreage diversity effects on both yield and chemical input use levels. These effects are assessed on a crop per crop basis, which allows to consider precisely defined agronomic effects and to circumvent the composition effects uncovered by Groom and Pereira Fontes (2021) for instance when considering crop aggregates. Furthermore, by considering pre crop effects, our modelling framework differs from that of Bareille and Letort (2018)[4] and explicitly accounts for dynamic features of crop production underlying crop rotation effects, contrary to Chavas and Kim (2007 and 2010), Chavas (2009) and Chavas and Di Falco (2012).

Our approach is based on a rich dataset that we organized by matching crop sequence acreages, obtained from administrative data, and soil and climate data to a cost accounting panel dataset covering 769 farms located in the Marne *département* (which is a small administrative district located around 150 km east from Paris) and its neighbourhood from 2008 to 2014. This particular area is of special interest because it counts among the most productive and the most diversified arable crop production areas in Europe. The Marne *département* provides local outlets to major grain crops, but also to sugar beet, potatoes and alfalfa.

We also developed purposely designed micro-econometric yield and input use models. These models feature pre crop and crop acreage diversity effects, thereby allowing the

[4] Our study also differs from that of Bareille and Letort (2018) in that we consider an extended version of their crop acreage diversity indicator. Our indicator proves to be empirically useful when the number of grown crops vary significantly across observation units.

investigation of important (mostly dynamic) features of crop production. Our simultaneously considering crop yield functions together with chemical input demand models enable us to estimate the effects of pre crops and crop acreage diversity on both crop yield and chemical input use levels. Our models also incorporate random parameters for accounting for farm and farmer heterogeneity in crop yield and input use levels, and in input productivity levels (*e.g.*, Suri, 2011; Koutchadé *et al*, 2018). Importantly, the rich information content of our dataset and our considering random parameter systems combining crop yield functions and input demand models allows controlling for the well-known input use endogeneity issues that arise when considering the estimation of production functions (*e.g.*, Griliches and Mairesse, 1995; Ackerberg *et al*, 2015).

The contributions of this article are threefold. First, we provide a modelling framework that allows identifying pre crop and crop acreage diversity effects on both yield and chemical input use levels, while accounting for the effects of farms' and farmers' unobserved heterogeneity and controlling for potential input use endogeneity in the considered econometric yield functions. Second, we obtain original empirical results on the effects of crop diversification on crop production in commercial farms. Our estimation results notably demonstrate statistically significant, albeit economically limited, pre crop and crop acreage diversity effects on crop yield levels and crop acreage diversity effects on herbicide use levels. Third, we obtain original empirical results on farmer choices regarding the effects of crop diversification. Our crop sequence acreage data demonstrate that farmers basically choose the best available pre crops for the major crops of their crop. Their choices are thus rational from both agronomic and economic viewpoints. This however reduces the scope of the crop diversification effects that can be assessed based on farm data.

Taken together our results show that pre crop effects can be uncovered from farm data, but only for the most frequent crop sequences. They also suggest that pre crop and crop acreage diversity effects provide insufficient economic benefits to farmers, especially insufficient savings of chemical inputs, for leading them to adopt diversified crop acreages.

The rest of the article is organized as follows. First, we present the models we use for uncovering pre crop and crop diversity effects from cost accounting data. Then, we describe our estimation strategy, with special emphasis on the issues raised by input use endogeneity and random parameters. The fourth section presents the different datasets we use and how we combine them. The fifth section presents and discusses the estimation results we obtain while the last one provide concluding remarks.

2. Modelling framework

Our primary interest lies in the magnitude of two types of effects on crop yield and input use levels, that is to say the effects of previous crops of the crop grown on the same plot (pre crop effects) on the one hand and the effects of crop acreage diversity at the farm level on the other hand. We consider estimating these effects based on a panel data set describing the production choices and performances of a large sample of farmers, $i=1,...,N$, over a short time period, $t=1,...,T$. Term $\mathcal{K}$ denotes the set of crops potentially grown by the sampled farmers, with $\mathcal{K}=\{1,...,K\}$.

The yield level of crop *k* obtained by farmer *i* in year *t* is denoted by $y_{k,it}$. The corresponding use level of input *j* is denoted by $x_{j,k,it}$ for $j\in\mathcal{J}$. In our application, the input set $\mathcal{J}=\{1,...,J\}$ includes nitrogen fertilizers, herbicides and other pesticides, which mostly include fungicides and insecticides.

Pre crop effects imply that the observed yield level of crop *k*, $y_{k,it}$, is a weighted average of the yield levels obtained for each previous crop after which crop *k* was grown by farmer *i* in year *t*. Let $y_{mk,it}$ denote the yield level of crop *k* when this crop is grown after crop *m* and let $z_{mk,it}$ denote the share of acreage of crop *k* grown after crop *m*. The observed yield level of crop *k* is given by

(1a) $y_{k,it} = \sum_{m \in \mathcal{K}} z_{mk,it} y_{mk,it}$ for $k \in \mathcal{K}$.

Similarly, accounting for potential pre crop effects on input use levels supposes to define input uses at the crop sequence levels. Let $x_{j,mk,it}$ denote the quantity of input *j* used by farmer *i* in year *t* for crop *k* when this crop is grown after crop *m*. The observed use level of input *j* for crop *k* is given by

(1b) $x_{j,k,it} = \sum_{m \in \mathcal{K}} z_{mk,it} x_{j,mk,it}$ for $k \in \mathcal{K}$ and $j \in \mathcal{J}$.

Crop yields and input uses being unobserved at the crop sequence level, we replace them by relatively simple models.

2.1. Yield and input use models at the crop sequence level

The model of yield at the crop sequence level is defined by

(2a) $y_{mk,it} = a_{mk,0}^{(y)} + \mu_{k,i}^{(y)} + \alpha_{k,t,0}^{(y)} + \mathbf{x}'_{mk,it}\boldsymbol{\beta}_{k,i} + \mathbf{c}'_{it}\boldsymbol{\lambda}_{k,0}^{(y)} + \mathbf{d}'_{it}\boldsymbol{\delta}_{k,0}^{(y)} + \varepsilon_{k,it}^{(y)}$, for $m \in \mathcal{K}$ and $k \in \mathcal{K}$,

where vector $\mathbf{x}_{mk,it} = (x_{j,mk,it}, j \in \mathcal{J})$ collects input uses at the crop sequence level. These input uses are included in the yield models for capturing the effects of the intensity in chemical inputs of farmers' production practices. These effects, represented by marginal productivity parameter $\boldsymbol{\beta}_{k,i}$, are assumed to be farm-specific. We consider crop yield functions instead of crop supply functions for disentangling the effects of pre crops or of crop acreage diversity that directly impact yield levels from those that may impact yield levels through adjustments in input use levels.

Vector $\mathbf{c}_{it}$ contains control variables aimed to capture the effects of production conditions that impact farmers' crop production choices. We use a rich set of variables describing soil properties at the farm level and weather conditions at the municipality level.

Farm specific parameters $\mu_{k,i}^{(y)}$ and $\boldsymbol{\beta}_{k,i}$ are expected to significantly vary across farmers and farms. Indeed, models featuring random parameters allow accounting for unobserved heterogeneity (*e.g.*, Wooldridge, 2010; Arellano and Bonhomme, 2011). In the case of agricultural production choices, this heterogeneity is due to unobserved characteristics of the sampled farmers (*e.g.*, skills, motivations) or farms (*e.g.*, spatial distribution of the plot, available machinery, unobserved soil or climate features) that do not vary or vary little over the considered time period (*e.g.*, Koutchadé *et al*, 2018 and 2021).

Year specific effects $\alpha_{k,t,0}^{(y)}$ capture the effects of large-scale factors impacting all farms, such as weather driven pest and disease outbreaks or widely adopted technological or agronomic innovations.

Parameter $a_{mk,0}^{(y)}$ defines the pre crop effects of crop m on the yield of crop k. Although these effects may vary across farms and years, we specify these effects as fixed parameters due to data constraints. Indeed, as discussed below, crop sequence acreages reported in farm datasets are likely to lack the variability needed to uncover year specific or the distribution of farm specific pre crop effects.

Vector $\mathbf{d}_{it}$ contains a set of crop acreage diversity indicators aimed to capture crop acreage diversity effects. These indicators combine grown crop numbers and Shannon indices. More precisely, vector $\mathbf{d}_{it}$ includes two subsets of variables: *(i)* a set of dummy variables, one for each number of crops grown observed in our data, for capturing the effect of the size of the crop set considered by farmers ; *(ii)* the cross products of these dummy variables with the corresponding crop acreage Shannon indices, for capturing the effects of the land allocation chosen by farmers.[5] Using this set of crop acreage diversity indicators has two main advantages. This allows estimating the effects of the number of grown crops without any parametric restriction on the one hand, and this allows circumventing a shortcoming of the Shannon index,

[5] Shannon indices are centered at their sample means by crop number for facilitating the interpretation of the direct effects of the crop number indicators.

explained in what follows, when the number of crops is variable on the other hand.

Let $s_{k,it}$ denote the share of the acreage of crop *k* in the arable land area of farmer *i* in year *t*. Provided that $s_{k,it} \in [0,1]$ for $k \in \mathcal{K}$ and $\sum_{k\in\mathcal{K}} s_{k,it} = 1$, the Shannon index of crop acreage $\mathbf{s}_{it} = (s_{k,it}, k \in \mathcal{K})$ is defined by $h(\mathbf{s}_{it}) = -\mathbf{s}_{it}' \ln \mathbf{s}_{it}$.[6] Function $h(\mathbf{s}_{it})$ is upper bounded by $\ln n(\mathbf{s}_{it})$ where $n(\mathbf{s}_{it}) \in \mathcal{K}$ is the number of crops actually grown in $\mathbf{s}_{it}$.[7] This upper bound grows at a rate that decreases in the number of crops actually grown, implying that the Shannon index is a questionable measure of crop diversity when numbers of grown crops vary widely across farms. Including in $\mathbf{d}_{it}$ the dummy variables corresponding to grown crop numbers and their cross products with Shannon indices allows disentangling the effects of grown crop numbers and those of the crop acreage diversity given grown crop numbers.

The models of input use at the crop sequence level are defined similarly, with

(2b) $\quad x_{j,mk,it} = a_{j,mk,0}^{(x)} + \mu_{j,k,i}^{(x)} + \alpha_{j,k,t,0}^{(x)} + \mathbf{d}_{it}'\boldsymbol{\delta}_{j,k,0}^{(x)} + \mathbf{c}_{it}'\boldsymbol{\lambda}_{j,k,0}^{(x)} + \varepsilon_{j,k,it}^{(x)}$ for $m \in \mathcal{K}$, $k \in \mathcal{K}$ and $j \in \mathcal{J}$

Parameter $a_{j,mk,0}^{(x)}$ is the pre crop effects of crop *m* on the use of input *j* for crop *k* and random term $\mu_{j,k,i}^{(x)}$ is a standard additively separable farm specific effect.[8] Year specific effects $\alpha_{j,k,t,0}^{(x)}$ capture the effects of large-scale factors, such as weather driven pest and disease outbreaks, widely adopted innovations or economic factors, including prices or changes in the value chains.

Due to the limited time span of our dataset, the effects of prices are difficult to disentangle from those of unobserved (to the analyst) temporal shocks or trends that impact all farmers. The

[6] Given that $s_{k,it} \ln s_{k,it}$ can be set at 0 if $s_{k,it} = 0$ by continuity extension, following the continuity of function $g(x) = x \ln x$ in $x \in \mathbb{R}_+^*$ and $\lim_{x\to 0^+} g(x) = 0$.

[7] Entropy function $h(\mathbf{s})$ achieves its (unique) maximum in $\mathbf{s} = (s_1, ..., s_K)$ at $h(\mathbf{s}) = \ln K$ and $h(\mathbf{s}) = \ln K$ if and only if $s_k = 1/K$ for $k \in \mathcal{K}$.

[8] Pre crop effect parameters $a_{mk,0}^{(y)}$ and $a_{j,mk,0}^{(x)}$ need to be normalized. They are set at 0 for a reference pre crop of each crop.

effects on chemical uses of crop and input prices, which mostly vary across years,[9] are expected to be captured by year specific terms $\alpha_{j,k,t,0}^{(x)}$. These price patterns also significantly impact our strategy for identifying the parameters of crop yield models (2a) since they basically prevent us from using prices as instrumental variables for input uses in the yield function equations we consider.

2.2. Yield and input use models at the crop level

Combining equations (1) and (2) yields crop *k* yield model

(3a) $$y_{k,it} = \mu_{k,i}^{(y)} + \alpha_{k,t,0}^{(y)} + \mathbf{x}_{k,it}'\boldsymbol{\beta}_{k,i} + \mathbf{z}_{k,it}'\mathbf{a}_{k,0}^{(y)} + \mathbf{d}_{it}'\boldsymbol{\delta}_{k,0}^{(y)} + \mathbf{c}_{it}'\boldsymbol{\lambda}_{k,0}^{(y)} + \varepsilon_{k,it}^{(y)}$$

and the corresponding input use models

(3b) $$x_{j,k,it} = \mu_{j,k,i}^{(x)} + \alpha_{j,k,t,0}^{(x)} + \mathbf{z}_{k,it}'\mathbf{a}_{j,k,0}^{(x)} + \mathbf{d}_{it}'\boldsymbol{\delta}_{j,k,0}^{(x)} + \mathbf{c}_{it}'\boldsymbol{\lambda}_{j,k,0}^{(x)} + \varepsilon_{j,k,it}^{(x)} \text{ for } j \in \mathcal{J}$$

for $k \in \mathcal{K}$. We use here the fact that the elements of vector $\mathbf{z}_{k,it}$, that is to say the shares of crop *k* on its potential pre crops, sum to 1.

The yield function models given in equation system (3) are linear in input use vector $\mathbf{x}_{k,it}$, which is unusual and, thereby, deserves a few comments. Admittedly, the linearity of crop sequence yield models (2a) in crop sequence input use levels $\mathbf{x}_{mk,it}$ greatly facilitates the aggregation process of these models at the crop level as shown by equation (3a). Yet, linear crop sequence yield models can be interpreted as first order Taylor expansion in $\mathbf{x}_{mk,it}$ of any (sufficiently smooth) model of $y_{mk,it}$. Importantly, the intercept and the coefficients of $\mathbf{x}_{mk,it}$ in the model of $y_{mk,it}$, which are collected in vector $\boldsymbol{\beta}_{k,i}$, are farm specific. Yield functions considered in equations (2a) and (3a) can thus be interpreted as farm specific approximates of the underlying "true" yield functions. Also, coefficient $\beta_{j,k,i}$ delivers a direct measure of the marginal

[9] The intra-farm variance of output prices does not exceed 20% of the total variance of crop prices in our sample. In our application, input prices are measured by price indices, which only vary in the time dimension.

productivity of input j for crop k in farm i.

3. Identification and estimation

Equations (3) define the production choice equations that we consider for estimation purpose. These are collected in the following equation system:

$$(4)\qquad \begin{cases} y_{k,it} = \mu_{k,i}^{(y)} + \alpha_{k,t,0}^{(y)} + \mathbf{z}_{k,it}'\mathbf{a}_{k,0}^{(y)} + \mathbf{x}_{k,it}'\boldsymbol{\beta}_{k,i} + \mathbf{d}_{it}'\boldsymbol{\delta}_{k,0}^{(y)} + \mathbf{c}_{k,it}'\boldsymbol{\lambda}_{k,0}^{(y)} + \varepsilon_{k,it}^{(y)} \\ \mathbf{x}_{k,it} = \boldsymbol{\mu}_{k,i}^{(x)} + \boldsymbol{\alpha}_{k,t,0}^{(x)} + \mathbf{Z}_{it}\mathbf{a}_{k,0}^{(x)} + \mathbf{D}_{it}\boldsymbol{\delta}_{k,0}^{(x)} + \mathbf{C}_{k,it}\boldsymbol{\lambda}_{k,0}^{(x)} + \boldsymbol{\varepsilon}_{k,it}^{(x)} \end{cases}.$$

Vectors $\boldsymbol{\mu}_{k,i}^{(x)}$, $\boldsymbol{\alpha}_{k,t,0}^{(x)}$ and $\boldsymbol{\varepsilon}_{k,it}^{(x)}$ are given by $\boldsymbol{\mu}_{k,i}^{(x)} = (\mu_{j,k,i}^{(x)} : j \in \mathcal{J})$, $\boldsymbol{\alpha}_{k,t,0}^{(x)} = (\alpha_{j,k,t,0}^{(x)} : j \in \mathcal{J})$ and $\boldsymbol{\varepsilon}_{k,it}^{(x)} = (\varepsilon_{j,k,it}^{(x)} : j \in \mathcal{J})$ while matrices $\mathbf{Z}_{it}$, $\mathbf{D}_{it}$ and $\mathbf{C}_{k,it}$ are defined by $\mathbf{Z}_{it} = \boldsymbol{\iota}_J \otimes \mathbf{z}_{it}'$, $\mathbf{D}_{it} = \boldsymbol{\iota}_J \otimes \mathbf{d}_{it}'$ and $\mathbf{C}_{k,it} = \boldsymbol{\iota}_J \otimes \mathbf{c}_{k,it}'$.

Estimating equation system (4) requires identification assumptions. Let vector $\boldsymbol{\gamma}_{k,i}$ collect the random parameters of model (4), with $\boldsymbol{\gamma}_{k,i} = (\boldsymbol{\mu}_{k,i}, \boldsymbol{\beta}_{k,i})$ and $\boldsymbol{\mu}_{k,i} = (\mu_{k,i}^{(y)}, \boldsymbol{\mu}_{k,i}^{(x)})$ and let vector $\boldsymbol{\varepsilon}_{k,it} = (\varepsilon_{k,it}^{(y)}, \boldsymbol{\varepsilon}_{k,it}^{(x)})$ collect its error terms. We assume that vectors $\boldsymbol{\gamma}_{k,i}$, $\boldsymbol{\varepsilon}_{k,it}$ and $\mathbf{w}_{k,it} = (\mathbf{z}_{k,it}, \mathbf{d}_{it}, \mathbf{c}_{it})$ are mutually independent and that the regressors collected in $\mathbf{w}_{k,it}$ are strictly exogenous in the considered equation system. Finally, we assume that error terms $\boldsymbol{\varepsilon}_{k,it}$ are serially uncorrelated. This last assumption makes use of the fact that farm specific random parameters are expected to capture the most persistent (unobserved) features of the dynamics of the considered crop production processes. Assuming that error terms $\boldsymbol{\varepsilon}_{k,it}$ and random parameters $\boldsymbol{\gamma}_{k,i}$ are independent is standard, and required for identifying the probability distribution of $\boldsymbol{\gamma}_{k,i}$.

3.1. Identification

The exogeneity assumption related to crop sequence acreage share vector $\mathbf{z}_{\ell,it}$ deserves some comments. These variables describe choices of farmers. Exogeneity of $\mathbf{z}_{\ell,it}$ is partly supported by the fact that crop sequence acreage decisions are taken prior to the occurrence of most random events impacting crop yield and input use levels. Yet, omitted variable biases may still arise. For instance, soils of farms with relatively large acreage shares of potatoes and/or sugar beet are generally deep and well structured. These soil properties have positive impacts on the production of most arable crops (*e.g.*, Carpentier and Letort, 2014). In our models soil property effects are controlled for by a rich set of variables describing the soils of the sampled farms.[10]

Similar observations hold regarding crop acreage diversity indicator vector, $\mathbf{d}_{it}$. In our empirical application, this vector is evaluated by considering previous year crop acreage shares, that is to say based on $\mathbf{s}_{it-1}$. This eliminates potential endogeneity issues related to error terms $\boldsymbol{\varepsilon}_{k,it}$. Our results on the effects of crop acreage diversity are robust to alternative construction approaches for the diversity indicator set $\mathbf{d}_{it}$, that is to say based on current acreages, on whole (observed) crop acreage history of farmers or on one year lagged crop acreages. This reflects the fact that farmers' crop acreages are relatively stable over time. This also provides support to our interpreting our empirical results on the effects of $\mathbf{d}_{it}$ as cropping system effects, at least to some extent. These are long run effects that are induced by the crop acreage history of the considered farm, which in turn determines the state of the farm agroecosystem. Importantly, the variables describing soil properties prevent our crop diversity indicators to capture the effects of soil quality. As discussed above, good soils enlarge the set of profitable arable crops for farmers. Failing to control for soil quality is likely to bias our empirical measure of crop acreage diversity effects through crop acreage diversity indicators $\mathbf{d}_{it}$.

We impose parametric distributional assumptions on random vectors $\boldsymbol{\gamma}_{k,i}$ and $\boldsymbol{\varepsilon}_{k,it}$, mostly for facilitating the estimation of the fixed parameters of the models and of the probability distribution of $\boldsymbol{\gamma}_{k,i}$ by Maximum Likelihood (ML). We assume that $\boldsymbol{\gamma}_{k,i}$ is multivariate normal, with $\boldsymbol{\gamma}_{k,i} \sim \mathcal{N}(\boldsymbol{\eta}_{k,0}, \boldsymbol{\Omega}_{k,0})$, and that error terms $\varepsilon_{k,it}^{(y)}$ and $\boldsymbol{\varepsilon}_{k,it}^{(x)}$ are normal, with $\varepsilon_{k,it}^{y} \sim \mathcal{N}(0, \psi_{kk,0}^{(y)})$ and

[10] The control variables describing farm soils include measures of soil depth, cationic exchange capacity, pH, water holding capacity as well as organic matter, clay, silt and sand contents.

$\boldsymbol{\varepsilon}_{k,it}^{(x)} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\psi}_{k,0}^{(x)})$. The correlation structure of $\mu_{k,i}^{(y)}$, $\mu_{k,i}^{(x)}$ and $\boldsymbol{\beta}_{k,i}$ is thus defined by variance-covariance matrix $\boldsymbol{\Omega}_{k,0}$.

Our last assumption concerns the status of the input use vector, $\mathbf{x}_{k,it}$, in the yield function model. This point is crucial as the endogeneity of input use in production function is a longstanding issue that has originated a considerable, and still lively, econometric literature (*e.g.*, Mundlak, 1996 and 2001; Just and Pope, 2001; Ackerberg *et al*, 2015). We assume that error terms $\boldsymbol{\varepsilon}_{k,it}^{(x)}$ and $\varepsilon_{k,it}^{y}$ are uncorrelated, implying input choices are assumed exogenous in the yield models conditionally on farm specific parameter $\boldsymbol{\gamma}_{k,i}$. This exogeneity assumption is admittedly restrictive, though common for panel data production function models (*e.g.*, Blundell and Bond, 2000; Suri, 2011).

Assuming that input uses $\mathbf{x}_{k,it}$ are exogenous with respect to $\varepsilon_{k,it}^{y}$ appears reasonable in our application, due to our using a rich set of control variables and our specifying farm specific random parameters. We basically assume that control variables $\mathbf{w}_{k,it}$, year specific effects $\alpha_{k,t,0}^{(y)}$ and farm specific random parameters $\mu_{k,i}^{(y)}$ and $\boldsymbol{\beta}_{k,i}$ capture most of the effects of the factors that simultaneously impact yield and input use levels. Importantly, the variance-covariance matrix, $\boldsymbol{\Omega}_{k,0}$, of random parameter vectors $\boldsymbol{\gamma}_{k,i}$ is left unrestricted.[11] Accordingly, input use levels $\mathbf{x}_{k,it}$ can be correlated with the random parts of the corresponding yield model, albeit only through correlations of the farm specific parameters of the yield model, $\mu_{k,i}^{(y)}$ and $\boldsymbol{\beta}_{k,i}$, and those of the input use equations, $\boldsymbol{\mu}_{k,i}^{(x)}$. The considered yield models thus accommodate a rich set of input use endogeneity sources.

[11] Implying that the considered yield models are so-called correlated random coefficient (linear) models (*e.g.*, Wooldridge, 2005a and 2005b; Suri, 2011)

3.2. Estimation

Equation (4) describes a recursive simultaneous equation system since input uses $\mathbf{x}_{k,it}$ are used as explanatory variables of yield levels $y_{k,it}$ while $y_{k,it}$ is not part of the models of $\mathbf{x}_{k,it}$. The input use and yield models featured in equation (4) rely on identical sets of exogenous variables, year specific dummy variables included. As argued above, we do not identify the parameters of our production choice models by relying on instrumental variables. Our identification strategy relies on a full information approach. It combines the use of a rich set of control variables and the parametric specification of the multivariate probability distribution function of the random parameters of the considered models. Basically, the control variables and the random parameters of the model enable us to assume that error terms $\boldsymbol{\varepsilon}_{k,it}^{(x)}$ and $\varepsilon_{k,it}^{y}$ are uncorrelated conditionally on $\mathbf{w}_{k,it}$ and $\boldsymbol{\gamma}_{k,i}$. We manage the issues raised by the endogeneity of $\mathbf{x}_{k,it}$ in the model of $y_{k,it}$ by explicitly modelling (*i*) the endogenous variable vector $\mathbf{x}_{k,it}$ and (*ii*) the correlation structure linking the farm specific random terms of the model of $\mathbf{x}_{k,it}$, $\boldsymbol{\mu}_{k,i}^{(x)}$, on the one hand and those of the model of $y_{k,it}$, $\mu_{k,i}^{(y)}$ and $\boldsymbol{\beta}_{k,i}$, on the other hand.

Although unusual, this approach is suitable given our data and objectives. As input use models (3b) are standard random parameter models, the estimation issues we face are mostly due to correlated random coefficient of crop yield models (3a). Breitung and Salish (2021) and Woodridge (2019) recently proposed alternative approaches, not requiring any distribution assumption on the random terms of the considered equation, for estimating correlated random coefficient models. These approaches however primarily focus on the estimation of random parameter means whereas we are also interested in there variance and covariance.

Since our models are fully parametric, we consider ML estimators of the parameters of equation systems (4). Apart from their being relatively large equation systems, two issues have to be dealt with when estimating these systems.

First, our models feature standard fixed parameters but also random coefficients, the parametric probability distribution of which is to be estimated. We circumvent this issue by using an Expectation-Maximization (EM) type algorithm (Dempster *et al,* 1977; Wu, 1983) for computing the ML estimators of the parameters of our models. EM type algorithms are particularly well suited for estimating models featuring unobserved variables, of which random parameters are prominent examples (*e.g*., Lavielle, 2014). Importantly, the interaction terms $\mathbf{x}'_{k,it}\boldsymbol{\beta}_{k,i}$ featured in the yield functions and the random parameters appearing in the input use models imply that the equation system we consider involve products of random parameters, which raises non-trivial estimation issues. Relying on stochastic extensions to standard (deterministic) EM algorithms is necessary in such nonlinear settings. The conditional expectation that constitutes the E step of EM algorithm cannot be integrated either analytically or numerically and, thereby, requires simulation methods. To that aim, we rely on a Stochastic Approximation Expectation Maximization (SAEM) algorithm, which was proposed by Delyon *et al* (1999) and makes a more efficient use of simulations than competing stochastic EM type algorithms. Our SAEM algorithm is presented in Appendix A.

Second, attrition is the rule rather than the exception in microeconomic panel data. Attrition is not an issue when units are missing at random, that is to say, when the decision to drop out is not related to factors that are correlated with the response variables (*e.g*., Wooldridge 2010). Our panel dataset is not balanced, for two reasons. First, farmers enter in and leave the customer base of the accountancy firm that made these data available to us. In addition, data can be lost or observations can be incompletely recorded. The resulting attrition processes can be considered as random regarding the modelled processes. Second, farmers can decide not to produce some crops. In our application, most sampled farmers produce the considered crops every year (wheat, barley or rapeseed) or do not produce them at all (sugar beet). Ignoring that observations are available depending on farmers' choice potentially raises endogenous sample selection issues when this choice depends on important unobserved factors that impact input use and yield levels. Yet, accounting for sample selection is relevant when the objective of the study is to infer the features of a process for an entire population while the process is only observed for a specific sub-population. The objective of our study is much simpler. We aim to estimate pre crop and crop acreage diversity effects and to investigate how farmers use them in the sample that is available to us. We do not seek for results that can be extrapolated to situations in which the considered crops could be produced whereas they are not produced in our data.

4. Data

Our dataset is primarily based on a large sample of farm cost accounting data and combines information from different sources for supplementing these basic data. We use data recorded by the IACS (Integrated Administration and Control System) in France for uncovering the crop sequence acreages of the sampled farms. We also use data from the *GlobalSoilMap* initiative (*e.g.*, Arrouays *et al* 2020) for obtaining measures of the main characteristics of the soils of the sampled farms. Finally, we used data from Meteo France for obtaining detailed information on the weather conditions that prevailed in the considered area over the considered time span.

4.1. Farm cost accounting data

Our main dataset consists of an unbalanced panel accountancy dataset of 769 farms mostly located in the Marne *département*, which provides a detailed description of farmers' choices in terms of acreages, crop yields as well as fertilizers and pesticides expenditures per crop. Aggregate fertilizers and pesticides volumes are computed from the corresponding expenditures by using price indices provided by the French ministry of agriculture at the regional level and expressed in constant 2010 euros per ha. Importantly, uses of the major nutrient elements – namely N, P and K – are also reported in kg for each crop.

We consider farms observed from 2008 to 2014 and displaying at least three consecutive years of observations. These farms mostly grow eight crops: wheat, barley, rapeseed, corn, protein pea, alfalfa, sugar beet and potatoes. We aim at estimating the pre crop effects of these eight crops on the yield and input uses of the four major crops of the sample: wheat, barley, rapeseed and sugar beet. Tables 1 reports some descriptive statistics on farmers' production choices for these four crops and on farms' acreage diversification. Wheat appears to be the dominant crop in our sample, as it is grown every year by all farms and represents on average one-third of farms' acreage. As shown by the figures reported in Table 1b, farms' acreages are quite diversified, with an average of five crops grown each year and most farmers (92% of them) growing at least four crops.

Table 1a. Descriptive statistics: yield and chemical input use levels

	Crops							
	Wheat		Barley		Rapeseed		Sugar beet	
Average yield (ton/ha)	8.65	(1.06)	7.00	(1.19)	3.88	(0.65)	93.02	(13.15)
Average output price (€/ton)	160	(33)	164	(35)	357	(71)	26	(4)
Average use of nitrogen (kg/ha)	217	(34)	147	(25)	214	(38)	137	(36)
Average use of herbicides (€/ha)	63	(19)	30	(12)	99	(31)	160	(53)
Average use of other pesticides (€/ha)	125	(36)	76	(25)	109	(38)	96	(31)
Average acreage share	0.34	(0.10)	0.21	(0.10)	0.14	(0.08)	0.12	(0.08)

Note. Sample standard deviation are in parentheses.

Table 1b. Descriptive statistics: crop acreage diversity

	Sample share	Average number of grown crops		Average Shannon index	
Farms growing 3 or less crops	0.08	2.85	(0.18)	1.09	(0.18)
Farms growing 4 crops	0.17	4.00	(-)	1.31	(0.10)
Farms growing 5 crops	0.37	5.00	(-)	1.48	(0.09)
Farms growing 6 crops	0.30	6.00	(-)	1.63	(0.08)
Farms growing 7 or more crops	0.08	7.04	(0.04)	1.76	(0.08)
Total sample		5.33	(1.02)	1.49	(0.21)

Note. Sample standard deviation are in parentheses.

4.2. IACS data

The IACS ensures the management of agricultural payments across EU countries[12]. It consists of several interconnected databases, which are updated on a yearly basis. We focus here on one particular database, the LPIS (Land Parcel Identification System). This database provides detailed information on crop acreages implemented on blocks of plots. Each plot block is geolocalized, has a unique identifier and is associated to a unique farm identifier. This information was used for uncovering the crop sequence acreages of the farms of our sample. The LPIS dataset is very rich but, because farm and plot block identifiers vary from one year to the other and because each block may contain more than one plot each year, specific data processing is required to extract information on farm crop sequences. For this purpose, we used the *RPG explorer* software developed by Martin *et al* (2017). By relying on established rules on how the crops match up from one year to the next, this software offers the possibility to recover crop sequence acreages shares at the farm level.[13] The extracted data were matched to our cost accounting data by using the farm acreage histories observed in both datasets as matching criteria. This allowed a perfect match[14] of two-thirds of the farms of the cost accounting data sample with a farm in the IACS data. More flexible matching criteria, combined with manual checks, then resulted in a matching rate of 78% of our initial sample.[15]

[12] More details can be found in European Commission (2019).

[13] These rules are based on the knowledge of the crop acreages declared in each block from one year to the next. A succession of 10 rules implemented one after the other can be used to determine the crop sequence. Rule 1 (only one crop grown in the block in one year and the next) is considered as giving certain results, rule 2 (two crops per year, distributed over equal areas from one year to the next but different from each other) and 3 (one crop in one year is "broken down" into two crops in the next year with equal total acreage) are supported by solid hypotheses, from rule 4 onwards the probability of error in identifying the sequence increases. In our case, it hasn't been necessary for the software to frequently make use of rules beyond rule 3. More details on each rule can be found in Martin *et al* (2017).

[14] We consider a match as perfect when, for each the eight selected crops, the discrepancy of the acreages between the two datasets is less than 0.1 hectare

[15] It should be noted here that, although very unlikely, a farm in our sample may have been matched with the wrong farm in the IACS sample. This could happen if the two farms are the same size and have very similar acreages each year. However, this would not have a significant

4.3. Soil and climate data

A significant part of the heterogeneity in yields and input uses among farms may be due to the heterogeneity in the soil and climatic conditions they face. To control for these heterogeneous factors in our econometric estimations, we also introduced soil and climate data in our dataset. We used weather indicators provided at the municipality level by *Meteo France,* the French National Weather Service, and soil quality indicators provided at the farm level[16] by the team of the "SoilServ" project funded by the French National Research Agency (ANR-16-CE32-0005). Statistics summing up these climate, respectively soil, indicators are reported in tables C1, respectively C2, of Appendix C.

4.4. Expert knowledge information

Expert knowledge information was gathered through interviews with three agricultural scientists and an extension agent of the considered area. These experts provided a list of unwarranted crop sequences, which are less likely to be chosen or never chosen by farmers and rankings of the effects of crop sequences on yield and input use levels. This information, which subsequently use to assess the consistency of our estimation results, is presented in Appendix D.

5. Results

Our application considers four crops: wheat, barley, rapeseed and sugar beet. These crops are

---

impact on our estimation results, since two farms with similar cropping patterns are expected to show similar effects of crop sequences and crop diversification on yields and input uses.

[16] Soil quality indicators were actually first provided at the plot level and aggregated at the farm level by using the correspondence between farms and plots provided by the IACS data.

the major crops in the considered area since they account for 82% of the arable crop acreage. We present here three sets of results. First, we briefly discuss the crop sequence acreages based on IACS data for our farm sample. Second, we present the estimation results of our equation systems composed of a yield equation and three inputs equations for nitrogen fertilizers, herbicides and other pesticides for each of the considered crops. We focus our discussion on the pre crop and crop acreage diversity effects. Third, we provide a set of results aimed to assess the economic value of the pre crop and crop acreage diversity effects that we uncover.

5.1. Crop sequence acreages

The average crop sequence acreage shares recovered from IACS data and matched to our data sample are reported in Table 2. These shares describe how farmers allocate the acreage of their current crops to plots with specific previous crops on average (*i.e.*, they correspond to the means of terms $z_{mk,it}$ over our sample). A striking feature of the results displayed in Table 2 is that farmers tend to rely on the most favourable crop sequences from an agronomic viewpoint. For instance, they avoid growing wheat after wheat (*i.e.*, the acreage share of wheat grown after wheat in the wheat acreage equals 0.06) and prefer growing wheat after rapeseed or sugar beet. The rapeseed-wheat sequence accounts for 39% of wheat crop acreage on average, the sugar beet-wheat sequence for 16%. We can also see that land previously used to grow pea, which is considered as a favourable pre crop for wheat by agronomists, is mostly devoted to wheat production.

Conversely, the average acreage shares of many crop sequences are almost null, implying that these crop sequences are almost never observed in our data. Many of these unobserved crop sequences are strongly unwarranted from an agronomic viewpoint and thus (almost) never used by farmers. For instance, growing rapeseed after rapeseed foster pest and disease issues. These first results suggest that farmers' crop sequence acreage choices are particularly rational from an agronomic viewpoint and, as a result, from an economic viewpoint.

Crop sequence acreage share choices are also subject to constraints, such as the limited availability of favourable pre crops for the major crops of farmers' acreages. Of course, available previous crop acreages depend on famers' previous acreage choices. For instance, wheat and barley account for about half of the arable crop acreage in the considered area. This implies that wheat and barley are major pre crops in the area, according to purely "mechanical" effects. Conversely, despite pea being among the most favourable previous crops for wheat and pea production being almost always followed by wheat production, the pea-wheat sequence only accounts for 6% of wheat crop acreage on average, because the average acreage share of pea only amounts to 0.02 in the area.

Farmers favouring the most profitable crop sequences and tending to avoid the least favourable ones significantly affects the econometric analysis of pre crop effects. First, the effects of pre crops involved in never or rarely used crop sequences are poorly identified, when they can be identified, because the corresponding crop sequence acreages display insufficient variability. Second, minor crops are generally grown on a few previous crops (which are often major crops[17]), if not on only one previous crop. Third, major crops are the only ones displaying allocations to pre crop acreages that significantly vary across farms. These observations sum up the main shortcoming of farm data (*i.e.*, of observed data as opposed to experimental data) for investigating the effects of crop sequences on crop yield and input use levels. Our estimation results on pre crop effects largely support the points made here.

[17] Straw cereals in our application.

Table 2. Average allocation of the acreages of the major crops to previous crop acreages

		Current (major) crops				Previous crop acreage shares
		Wheat	Barley	Rapeseed	Sugar beet	
Previous crops	**Wheat**	0.06	0.56	0.38	0.59	0.35
	Barley	0.02	0.13	0.59	0.32	0.22
	Rapeseed	0.39	0.03	0.00	0.01	0.14
	Corn	0.06	0.03	0.00	0.00	0.03
	Protein pea	0.06	0.00	0.00	0.00	0.02
	Alfalfa	0.08	0.01	0.00	0.00	0.08
	Sugar beet	0.16	0.18	0.01	0.02	0.12
	Potatoes	0.05	0.02	0.00	0.02	0.02
	Other pre crops	0.10	0.03	0.01	0.03	0.02
Current crop acreage shares		0.34	0.21	0.14	0.12	

5.2. Estimation results

Table 3 displays simulated R^2 measures for our estimated models. These measures correspond to standard R^2 measures applied to models in which farm specific parameters are replaced by their estimates based on the estimated models and farm specific observations (*e.g.*, Koutchadé *et al* 2018, 2021). These simulated R^2 demonstrate that our models provide a better fit to the yield data than to the chemical input use data. The estimated crop yield models explain from 63% to 70% of the observed yield variance. The pesticide use models explain from 50% to 65% of the observed herbicide and other pesticide use variances. The corresponding percentages drop down to 38% and 44% for the nitrogen fertilizer use models.

Table 3. Simulated R^2 measures for the estimated yield and input uses models

	Crops			
	Wheat	Barley	Rapeseed	Sugar beet
Simulated R^2				
Yield models	0.68	0.63	0.64	0.70
Fertilizer use models	0.38	0.42	0.42	0.44
Herbicide use models	0.55	0.50	0.55	0.67
Other pesticide use models	0.65	0.55	0.60	0.55
Number of observations	3,982	3,327	3,530	3,085
Number of farms	769	654	692	607

Due to space limitation we focus our presentation of the estimation results on the parameters characterizing the marginal productivities of chemical inputs at the farm level (*i.e.*, random parameters $\boldsymbol{\beta}_{k,i}$), and on our main parameters of interest: the parameters capturing the effects of crop sequences (*i.e.*, fixed parameters $\mathbf{a}_{k,0}^{(y)}$ and $\mathbf{a}_{k,0}^{(x)}$) and crop acreage diversity (*i.e.*, fixed parameters $\boldsymbol{\delta}_{k,0}^{(y)}$ and $\boldsymbol{\delta}_{k,0}^{(x)}$) on yield and input use levels. The other estimation results are reported in Appendix C.[18]

5.2.1. Marginal productivity and return levels

The farm specific coefficients of the input use variables in the crop yield models (*i.e.*, random parameters $\boldsymbol{\beta}_{k,i}$) give the marginal productivity of the considered chemical inputs at the farm level. The estimates of the parameters (means and standard deviations) characterizing their distribution are reported in Table 4.

The estimated the means of $\beta_{j,k,i}$, which yield the average marginal productivities are of both sign and small in absolute value for nitrogen fertilizers. Such results are common in the agricultural production economics literature, at least for conventional production practices. Responses of crop production to nitrogen uses are known to generally exhibit a plateau at high nitrogen use levels.[19] The estimated average marginal productivity levels are positive for pesticides, and small for herbicides. Farmers usually use herbicides for controlling weeds following a long run strategy (*e.g.*, Colbach and Cordeau 2018). Herbicides aim to control current weed populations as well as weed seed banks. Parameters $\beta_{j,k,i}$ capture the short run effects of input *j* on the yield level of crop *k*. They fail to capture their long run effects, which

[18] Tables C3-C6 mostly report estimation results of the effects of weather and soil property variables. Including these variables in our models appears to be significantly improve their estimation from a statistical viewpoint.

[19] See, *e.g.*, Tembo *et al* (2008) for a recent empirical analysis of this feature of the productivity of nitrogen fertilizers.

are expected to be positive for herbicides.

Table 4. Estimated means and standard deviation of inputs marginal productivities ($\beta_{j,k,i}$ parameters)

	Wheat		**Barley**		**Rapeseed**		**Sugar beet**	
Nitrogen								
mean (x100)	0.03	(0.03)	-0.06	(0.05)	0.08**	(0.01)	-0.40	(0.39)
std deviation (x100)	0.21**	(0.03)	0.30**	(0.06)	0.15**	(0.01)	1.95**	(0.45)
Herbicides								
mean (x100)	0.14**	(0.05)	0.02	(0.11)	0.04**	(0.02)	0.55	(0.32)
std deviation (x100)	0.38**	(0.07)	0.63**	(0.15)	0.22**	(0.02)	2.42**	(0.40)
Other pesticides								
mean (x100)	0.34**	(0.03)	0.27**	(0.06)	0.13**	(0.02)	0.80	(0.61)
std deviation (x100)	0.17**	(0.03)	0.37**	(0.07)	0.20**	(0.02)	5.39**	(0.80)

Note. Symbol "**", respectively "*", indicates that the parameter is tested non-null at the 5%, respectively 10%, level. Estimated standard deviations of the parameter estimators are in parentheses beside the parameter estimates.

Let assume that input *j* is purchased at price $p_{j,k}^{(x)}$ and crop *k* is sold at price $p_{k}^{(y)}$. According to our crop yield models, the marginal net return of crop *k* to input *j* is given by $p_{k}^{(y)}\beta_{j,k,i} - p_{j,k}^{(x)}$ for farm *i*. Since input prices or price indices are close to one in our application this expression provides the net return to the last Euro of input *j* used for crop *k*. Table 5 reports the sample means of the estimated marginal net returns of crops to chemical inputs at the sample mean prices. According to these results, chemical inputs, nitrogen fertilizers and herbicides in particular, are used above their respective expected profit maximizing levels by most farmers. These results suggest that the farmers of our sample seek for relatively high yield levels, which supposes to rely on relatively high crop protection and fertilization levels. Yet, assessing the extent of the revealed overuses of chemical inputs would require assessing the concavity properties of the considered crop yield functions, which is out of the scope of this study.[20]

[20] If these overuse levels are significant they are unlikely to be "massive". As reminded above, crop yields are known to exhibit a plateau at high nitrogen use levels. Similarly, the marginal productivity of pesticides is known to sharply decrease in pesticide use levels (*e.g.*, Frisvold 2019).

Table 5. Estimated means of crop marginal net returns to chemical inputs (at average price levels, 2008–2014)

	Nitrogen fertilizers		Herbicides		Other pesticides	
Wheat	-0.95	(0.05)	-0.78	(0.08)	-0.46	(0.05)
Barley	-1.10	(0.08)	-0.97	(0.18)	-0.56	(0.10)
Rapeseed	-0.71	(0.04)	-0.86	(0.07)	-0.54	(0.07)
Sugar beet	-1.10	(0.10)	-0.86	(0.08)	-0.79	(0.16)

Note. Estimated standard deviations of the parameter estimates are in parentheses

5.2.2. Pre crop effects on yield and input use levels

Table 6 reports the estimated effects of pre crops on the yield and input use levels of the considered crops compared to that of a reference pre crop. The reference pre crop of a given crop is its most frequent one in our data, which is rapeseed for wheat and wheat for the other crops. As evidenced by the relatively large standard deviations of estimated effects reported in Table 6, previous crop effects are poorly identified in our models[21], particularly on input uses. As discussed above, such results can partly be due to crop sequence acreage choices patterns. Since farmers tend to select the most favourable pre crop for their major crops and to choose similar crop acreages across years, which limits the variability in crop sequence acreages and prevents obtaining accurate estimates of the corresponding pre crop effects.[22] The implied limited v. For instance, the acreage share of the pea-barley sequence is almost null on average while the barley-barley one equals 0.13 (Table 2). The estimated standard errors of pre crop effects of pea on barley yield and input uses are at least fivefold those of barley (Table 6). This comes to illustrate our concerns regarding the identification of crop rotation effects for rarely

[21] Even if their signs are generally consistent with experts' views.

[22] Small pre crop effects in absolute value could also be due to the well-known attenuation biases induced by measurement errors (*e.g.*, Wooldridge, 2010) since crop sequence acreage shares are reconstructed in our data. Yet, attenuation biases are also associated to downward biased estimates of the standard deviation of the estimated effects (*e.g.*, Wooldridge, 2010). This suggests that our pre crop effect estimates are unlikely to be impacted by significant attenuation biases in our application.

used crop sequences. Unsurprisingly, the most accurate estimated pre crop effects are those of the wheat yield equation. They tend to show that growing wheat after rapeseed instead of after wheat increases expected yield levels by 0.32 t/ha while growing wheat after pea instead of after wheat increases expected yield levels by 0.52 t/ha (which amounts to 6% of the sample average wheat yield). These results are in line with experimental results (*e.g.*, Meynard *et al* 2013, Jeuffroy *et al* 2015, Preißel *et al* 2015).

Our results demonstrate very limited adjustments, if any, of farmers' chemical input uses to the pre crops of their crop acreages. In particular, we decided to focus on nitrogen uses instead of aggregated fertilizer uses for investigating the effects on nitrogen uses of legumes as pre crops. Our results tend to show that farmers do not reduce their nitrogen uses after legumes. These estimation results are consistent with the results obtained by Nave *et al* (2013) based on French farmers interviews. They are also consistent with our estimates of the marginal productivity of nitrogen fertilizers, at least to some extent. Farmers tend to overuse nitrogen fertilizers when they downplay legume nitrogen surpluses while deciding their uses of mineral nitrogen. Nitrogen surpluses are not lost, as they can induce higher yields, but total nitrogen applications exceed their economically optimal levels.

Three phenomena may underlie these results. First, data issues may prevent our obtaining accurate estimates of pre crop effects. As discussed above, our data may not contain sufficient information for uncovering pre crop effects on input uses. Moreover, recorded crop input use levels are likely to be less accurate than recorded yield levels. These data issues cannot be ruled out but they are unlikely to fully explain our low estimates of pre crop effects on input uses. As will be shown below, these data allow us to uncover relatively precise crop acreage diversity effects on pesticide uses.

Second, if farmers automatically reap off the benefits of pre crop effects on yield levels they need to be willing to adjust their chemical input uses for benefiting from the effects of pre crops on soil nutrient contents and/or on pest and weed pressures. Farmers are likely to be aware of nitrogen surpluses left by legumes or of the break effects of diversified crop sequences on pest and weed pressures but they may be reluctant to adjust their input uses because nitrogen surpluses and break crop effects on pest and weed pressures are random and assessing them is costly.

Third, pre crop effects on chemical inputs may be limited in our data. Indeed, our data enable us to uncover the effects of crop sequences involving pre crops with sufficiently variable acreages for a given crop. These pre crops, those of wheat excepted, may not induce contrasted effects either on yield levels or on chemical input uses. For instance, wheat and barley are by far the most frequent pre crops of sugar beet in our sample. Yet, these crops are both straw cereals sharing many agronomic features and are expected to have similar effects on sugar beet production by the experts we have consulted. Accordingly, wheat being the reference pre crop of sugar beet in our application, the absence of statistically significant pre crop effects of barley on sugar beet production is not particularly surprising. Finally, farmers' production practices may also induce limited pre crop and crop acreage diversity effects. Experimental results of Coulter *et al* (2011) suggest that high yielding cropping management practices attenuate the effects of crop diversity. Indeed, high yields require high nutrient loads and high yielding crop cultural techniques tend to enhance pest, weed and disease pressures.[23] These effects may swamp those of crop diversity.

[23] The most productive seeds are also often more susceptible to pests and diseases. High nitrogen fertilization levels increase the susceptibility of crops to diseases. Early sowing increases the exposure of crops to pests and diseases while dense uniform sowing may foster the occurrence and the severity of pest and disease outbreaks. Such effects were discussed and documented by Loyce and Meynard (1997) and Loyce *et al* (2008 and 2012) in the case of winter wheat.

Table 6. Estimated pre crop effects ($a_{k,0}^{(y)}$ and $a_{k,0}^{(x)}$ parameters)

Pre crop effects	Wheat		Barley		Rapeseed		Sugar beet	
	Yield models (t/ha)							
Wheat	-0.32**	(0.10)	0.00	(-)	0.00	(-)	0.00	(-)
Barley	-0.19	(0.16)	-0.16**	(0.07)	0.04	(0.03)	-0.25	(0.55)
Rapeseed	0.00	(-)	0.16	(0.20)			1.25	(3.03)
Protein pea	0.20**	(0.10)	0.17	(0.41)	0.05	(0.25)		
Alfalfa	-0.09	(0.08)	0.02	(0.35)			-8.51	(5.22)
Sugar beet	0.01	(0.06)	0.21**	(0.07)			-0.51	(1.69)
	Fertilizer use models (kg/ha)							
	Wheat		**Barley**		**Rapeseed**		**Sugar beet**	
Wheat	0.32	(3.79)	0.00	(-)	0.00	(-)	0.00	(-)
Barley	2.01	(6.32)	0.85	(1.91)	-1.66	(1.77)	-2.39	(1.91)
Rapeseed	0.00	(-)	-0.74	(4.53)			6.36	(7.93)
Protein pea	1.83	(3.37)	-5.19	(9.81)	-4.79	(18.50)		
Alfalfa	4.04	(3.27)	5.06	(8.08)			2.51	(13.00)
Sugar beet	2.01	(2.54)	-0.12	(1.58)			-5.73	(5.75)
	Herbicide use models (€/ha, 2010 prices)							
	Wheat		**Barley**		**Rapeseed**		**Sugar beet**	
Wheat	-0.11	(2.29)	0.00	(-)	0.00	(-)	0.00	(-)
Barley	1.37	(4.13)	1.94**	(0.95)	-0.57	(1.44)	-2.57	(2.52)
Rapeseed	0.00	(-)	0.62	(2.46)			7.78	(11.21)
Protein pea	1.55	(2.38)	-5.17	(5.78)	5.52	(18.90)		
Alfalfa	0.58	(2.18)	1.62	(3.74)			9.24	(34.09)
Sugar beet	-1.35	(1.58)	-1.68*	(0.97)			8.41	(8.50)
	Other pesticide use models (€/ha, 2010 prices)							
	Wheat		**Barley**		**Rapeseed**		**Sugar beet**	
Wheat	4.07	(4.08)	0.00	(-)	0.00	(-)	0.00	(-)
Barley	-15.63**	(6.55)	-0.50	(1.88)	1.10	(1.73)	-1.05	(1.55)
Rapeseed	0.00	(-)	-3.34	(4.93)			-2.23	(6.85)
Protein pea	-5.32	(3.52)	-5.64	(17.04)	-16.49	(20.27)		
Alfalfa	-2.05	(3.27)	8.48	(4.42)			-3.92	(13.05)
Sugar beet	-1.66	(2.50)	0.87	(1.91)			1.86	(5.90)

Note. Symbol "**", respectively "*", indicates that the parameter is significantly estimated at the 5%, respectively 10%, level. Estimated standard deviations of the parameter estimates are in parentheses.

5.2.3 Crop acreage diversity effects on yield and input use levels

Table 7 reports the effects of the number of grown crops, our main crop acreage diversity effects, on the outcomes of the considered crops. These effects are measured as differences with respect to effect of the reference number of grown crops, which is five in our empirical application. We also report the effects of the crop acreage Shannon index interacted with the grown crop number dummy variables. These effects measure the effects of crop acreage diversity holding fixed the number of crops actually grown.

Crop diversity is expected by agronomists (*e.g*., Meynard *et al* 2013, Duru *et al* 2015) to increase crop yield levels, to decrease pesticide use levels and, to a lesser extent, to decrease fertilizer use levels. Our estimation results tend to support these hypotheses. Yield levels increase and pesticide use levels decrease as the grown crop number and the related Shannon index increase. Yet, the estimated effects are non-null from a statistical viewpoint only for yield and herbicide use levels.

According to our results, wheat yield levels are 0.20 t/ha higher when wheat is grown in a 7 crop farm than in a 5 crop farm, and wheat yield levels are 0.39 t/ha higher when wheat is grown in a 7 crop farm than in a 3 crop farm. Moreover, crop acreage diversity as measured by the Shannon index also tends to increase wheat yield levels holding constant the number of grown crops. Other crops display similar patterns regarding crop acreage diversity effects. Such results are fully consistent with agronomists' views and experimental results (*e.g*., Lin 2011, Kremen and Miles 2012, Meynard *et al* 2013, Hufnagel *et al* 2020).

Our results also tend to demonstrate that crop acreage diversity significantly impacts herbicide uses, although with less accurate estimates than for crop yields. For instance, farms growing 5 crops use 4.7 €/ha (at the 2010 price levels) less herbicides on wheat than farms growing 3 crops, farms spending 63.1 €/ha on herbicides for wheat on average. Similarly, farms growing 5 crops use 11.1 €/ha less herbicides on rapeseed than farms growing 3 crops, farms spending 99.3 €/ha on herbicides for rapeseed on average. Overall, these results are consistent with agronomists' experimental results (*e.g*., Chikowo *et al*, 2009; Adeux *et al*, 2019; Sharma *et al*, 2021).

The estimated effects of the crop acreage Shannon index per grown crop number also suggest that crop acreage diversity tends to lower weed pressures. Yet, the related estimates are generally inaccurate and lack statistical significance. In the same vein, the estimated effects of the crop acreage diversity indicators on sugar beet herbicide use levels have expected signs and are relatively large but are estimated too inaccurately for being meaningful from a statistical viewpoint.

Table 7. Estimated crop acreage diversity effects

	Yield models (t/ha)							
Crop diversity effects	**Wheat**		**Barley**		**Rapeseed**		**Sugar beet**	
3 crops or less	-0.19**	(0.07)	-0.35**	(0.13)	0.01	(0.07)	-7.67	(4.39)
4 crops	-0.10*	(0.05)	-0.19**	(0.07)	-0.04	(0.04)	-1.08	(1.29)
5 crops	0.00	(-)	0.00	(-)	0.00	(-)	0.00	(-)
6 crops	0.06	(0.04)	0.02	(0.05)	0.06*	(0.03)	1.38**	(0.61)
7 crops or more	0.20*	(0.07)	0.15*	(0.08)	0.18**	(0.04)	1.90*	(0.93)
3 crops x ***Shannon ind***	-0.13	(0.41)	0.69	(0.62)	-0.26	(0.36)	-13.27	(29.80)
4 crops x ***Shannon ind***	0.72**	(0.27)	1.33**	(0.47)	0.21	(0.20)	6.57	(6.89)
5 crops x ***Shannon ind***	0.73**	(0.31)	0.88**	(0.36)	0.33	(0.20)	9.79**	(4.58)
6 crops x ***Shannon ind***	0.28	(0.33)	0.40	(0.41)	0.59**	(0.23)	14.38**	(4.42)
7 crops x ***Shannon ind***	1.44**	(0.58)	0.87	(0.64)	0.93**	(0.30)	15.73**	(6.48)

	Fertilizer use models (kg/ha)							
	Wheat		**Barley**		**Rapeseed**		**Sugar beet**	
3 crops or less	-0.38	(3.20)	-1.41	(4.31)	-2.01	(4.82)	6.08	(24.09)
4 crops	0.40	(1.64)	0.69	(1.63)	-1.59	(2.28)	3.27	(4.41)
5 crops	0.00	(-)	0.00	(-)	0.00	(-)	0.00	(-)
6 crops	1.55	(1.40)	0.54	(1.12)	3.45*	(1.68)	-0.44	(1.91)
7 crops or more	2.65	(2.19)	1.37	(2.04)	3.72	(2.48)	-3.03	(2.60)
3 crops x ***Shannon ind***	-3.74	(17.91)	-13.71	(26.74)	-12.25	(19.33)	-83.64	(107.95)
4 crops x ***Shannon ind***	10.54	(9.13)	-6.55	(10.83)	-3.94	(14.56)	1.24	(24.56)
5 crops x ***Shannon ind***	-12.98	(11.09)	-12.00	(8.20)	-13.53	(13.49)	-36.25**	(16.84)
6 crops x ***Shannon ind***	-1.56	(11.34)	-4.22	(8.88)	12.62	(12.46)	1.78	(13.29)
7 crops x ***Shannon ind***	-3.28	(20.36)	-7.39	(17.68)	0.22	(22.09)	-25.68	(21.83)

	Herbicide use models (€/ha, 2010 prices)							
	Wheat		**Barley**		**Rapeseed**		**Sugar beet**	
3 crops or less	4.66**	(1.54)	3.56	(2.02)	11.05**	(3.63)	12.11	(59.26)
4 crops	1.93	(1.03)	1.20	(0.84)	8.14**	(1.88)	4.21	(3.53)
5 crops	0.00	(-)	0.00	(-)	0.00	(-)	0.00	(-)
6 crops	-1.64*	(0.88)	-1.37**	(0.58)	-2.42	(1.49)	-1.41	(2.39)
7 crops or more	-2.37	(1.74)	-1.34	(0.99)	-3.29	(2.48)	-5.30	(4.27)
3 crops x ***Shannon ind***	1.99	(8.68)	-12.01	(11.14)	15.37	(23.84)	-38.16	(372.78)
4 crops x ***Shannon ind***	-9.35	(5.15)	-10.32**	(4.98)	-8.71	(11.66)	-31.32	(18.48)
5 crops x ***Shannon ind***	-4.12	(6.13)	-8.88**	(3.68)	-12.98	(10.83)	-4.11	(19.37)
6 crops x ***Shannon ind***	-7.27	(8.17)	-3.05	(5.66)	-23.70*	(12.69)	3.41	(18.70)
7 crops x ***Shannon ind***	7.57	(15.25)	-0.67	(9.88)	-18.79	(20.45)	-11.79	(30.71)

	Other pesticide use models (€/ha, 2010 prices)							
	Wheat		**Barley**		**Rapeseed**		**Sugar beet**	
3 crops or less	-3.73	(3.28)	-2.44	(3.33)	-0.82	(4.15)	3.11	(18.21)
4 crops	-3.17	(2.07)	0.14	(1.66)	2.94	(2.37)	0.92	(2.53)

5 crops	0.00	(-)	0.00	(-)	0.00	(-)	0.00	(-)
6 crops	-0.25	(1.76)	-0.72	(1.32)	-1.35	(1.89)	1.00	(1.62)
7 crops or more	1.90	(2.89)	-2.16	(2.22)	0.75	(3.10)	-1.50	(2.79)
3 crops x ***Shannon ind***	5.99	(14.46)	1.52	(25.08)	-16.66	(28.23)	55.79	(73.47)
4 crops x ***Shannon ind***	1.83	(10.94)	-1.28	(8.37)	-1.55	(12.14)	-10.12	(13.18)
5 crops x ***Shannon ind***	-13.95	(11.38)	-6.53	(8.91)	-29.88**	(13.92)	-26.92**	(12.69)
6 crops x ***Shannon ind***	1.58	(13.52)	-5.49	(11.55)	23.23	(14.91)	3.80	(12.42)
7 crops x ***Shannon ind***	5.31	(17.92)	-1.52	(17.93)	4.22	(19.47)	10.15	(21.19)

Note. Symbol “**”, respectively “*”, indicates that the parameter is significantly estimated at the 5%, respectively 10%, level. Estimated standard deviations of the parameter estimates are in parentheses.

In addition, the effects of crop acreage diversification on herbicide uses may not solely be due to agro-ecological weed regulation effects. For instance, in our sample most farms growing at most 3 crops do not produce sugar beet or potatoes, implying that comparing farms growing 3 crops (or less) and farms growing at least 4 crops largely consists of comparing farms not growing sugar beet to farms growing sugar beet or potatoes. As these root crops require high-level weed control, the effects of crop acreage diversity on herbicide use on grain crops may be due to carry-over effects of the chemical weed control implemented for protecting root crops.[24]

It is worth noting that our results on crop acreage diversity effects in the yield and herbicide equations are not driven by farm size, as the inclusion of the farm size among the control variables turns out to be statistically insignificant regardless the equation in which it is included. In addition, our results are not driven by soil quality effects. Our including a rich set of soil property measures in our models controls for the well-known fact that good soils both widen the scope of profitable crops (*e.g.*, root crops require deep and suitably structured soils) and enhance crop profitability.[25] Yet, the fact that farmers using more diversified crop acreages may be more skilled from a technical viewpoint cannot be ruled out.

[24] This comes to illustrate the point recently made by Colbach *et al* (2020) on the analysis of farm versus experimental data when considering crop rotations, weed populations and their effects on crop yields, and herbicide uses.

[25] Notwithstanding our data covering a limited geographical area.

5.3. Economic assessment of pre crop and crop acreage diversity effects

In order to assess the economic value of the pre crop and crop acreage diversity effects uncovered by our modelling framework, we compute the effects of changes in pre crops or in crop acreage diversity on farmers' crop returns to chemical inputs. We compute these effects for an "average farmer" and at the mean prices of our sample. Appendix B provides the related technical details.

Table 8a reports the effects of changing the pre crop of wheat and barley from their respective reference pre crop to pre crops that have significant effects according to our estimation results. Growing wheat after wheat rather than after rapeseed entail an average loss of 53 €/ha, which amounts to 4.6% of the average wheat return to chemical inputs. On the contrary, growing wheat after protein pea rather than after rapeseed increases wheat crop return by 33 €/ha on average. Most of these effects on wheat return are due to pre crop effects on wheat yield levels, as pre crop effects on input uses are limited in general.

Changing the pre crop of barley to corn or barley entails losses while changing to sugar beet implies gains of around 35€/ha. As in the case of wheat, these effects on barley return are mostly due to pre crop effects on barley yield levels. Their magnitude is limited as they represent around 4% of the average barley return to chemical inputs.

Table 8a. Economic assessment of selected pre crop effects (€/ha), average effects at average price levels (2008–2014)

	Economic value of the effect of a pre crop, *versus* the considered crop reference pre crop (€/ha)									
	Yield value		**Nitrogen fertilizer cost**		**Herbicide cost**		**Other pesticide cost**		**Return to chemical inputs**	
	(1)		(2)		(3)		(4)		(1) - (2) - (3) - (4)	
	Wheat (reference pre crop: rapeseed)									
Wheat	-48.9**	(15.9)	0.1	(1.7)	-0.1	(2.3)	4.1	(4.1)	-53.1**	(16.0)
Protein pea	30.0*	(17.2)	0.8	(1.6)	1.5	(2.4)	-5.3	(3.5)	32.9*	(16.8)
Yield value or input cost sample mean	1,377		240		63		125		949	
	Barley (reference pre crop: wheat)									
Barley	-27.1**	(11.5)	0.6	(1.3)	1.9**	(0.9)	-0.5	(1.9)	-29.1**	(11.9)
Corn	-46.6**	(20.1)	-0.1	(2.8)	-2.1	(1.5)	-7.4**	(3.3)	-37.0	(20.5)
Sugar beet	34.6**	(11.5)	-0.1	(1.1)	-1.7	(1.0)	0.9	(1.9)	35.5**	(11.8)
Yield value or input cost sample mean	1,141		181		30		76		854	

Note. Symbol "**", respectively "*", indicates that the parameter is significantly estimated at the 5%, respectively 10%, level. Estimated standard deviations of the parameter estimates are in parentheses.

Table 8b reports the average effects of crop acreage diversity, as measured here by the number of grown crops, on crop returns to chemical inputs. Farmers growing wheat as part of a 7 crops acreage improves wheat return by 31.8 €/ha on average compared to growing wheat as part of a 5 crop acreage. Similarly, growing wheat as part of a 3 crops acreage instead of a 5 crop one entails an average loss of 32.1 €/ha.

Comparable results are obtained for the other considered crops. For instance, the return of barley is improved by 85.2 €/ha on average when the number of grown crops increases from 3 to 7. Increasing the number of grown crops from 5 to 7 leads to average return increases of 67.2 €/ha for rapeseed and of 57.7 €/ha for sugar beet.

Table 8b. Economic assessment of crop diversity effects (€/ha), average effects at average price levels (2008–2014)

	Economic value of the number of grown crops, *versus* 5 grown crops (€/ha)									
	Yield value		**Nitrogen fertilizer cost**		**Herbicide cost**		**Other pesticide cost**		**Return to chemical inputs**	
	(1)		(2)		(3)		(4)		(1) - (2) - (3) - (4)	
	Wheat (reference: 5 crops)									
3 crops or less	-31.4**	(12.1)	-0.2	(1.5)	4.7**	(1.5)	-3.7	(3.3)	-32.1**	(12.7)
4 crops	-17.8**	(8.5)	0.2	(0.8)	1.9*	(1.0)	-3.2	(2.1)	-16.8**	(8.7)
6 crops	9.2	(6.9)	0.7	(0.6)	-1.6*	(0.9)	-0.3	(1.8)	10.4	(6.7)
7 crops or more	32.5**	(11.3)	1.2	(1.0)	-2.4	(1.7)	1.9	(2.9)	31.8**	(11.7)
Sample mean	1,377		240		63		125		948	
	Barley (reference: 5 crops)									
3 crops or less	-58.1**	(21.0)	-1.0	(2.9)	3.6	(2.0)	-2.4	(3.3)	-58.4**	(21.2)
4 crops	-31.6**	(11.7)	0.5	(1.1)	1.2	(0.8)	0.1	(1.7)	-33.3**	(11.9)
6 crops	3.0	(8.4)	0.4	(0.8)	-1.4**	(0.6)	-0.7	(1.3)	4.8	(8.4)
7 crops or more	24.1*	(12.6)	0.9	(1.4)	-1.3	(1.0)	-2.2	(2.2)	26.8**	(12.5)
Sample mean	1,141		181		30		76		854	
	Rapeseed (reference: 5 crops)									
3 crops or less	4.5	(24.8)	-0.9	(2.2)	11.1**	(3.6)	-0.8	(4.2)	-4.6	(24.6)
4 crops	-10.4	(13.3)	-0.7	(1.1)	8.1**	(1.9)	2.9	(2.4)	-20.6	(13.4)
6 crops	21.9**	(9.8)	1.6**	(0.8)	-2.4	(1.5)	-1.3	(1.9)	23.6**	(9.6)
7 crops or more	66.9**	(15.9)	1.7	(1.2)	-3.3	(2.5)	0.8	(3.1)	67.2**	(15.6)
Sample mean	1,381		233		99		109		940	
	Sugar beet (reference: 5 crops)									
3 crops or less	-198.8	(104.3)	4.4	(17.4)	12.1	(59.3)	3.1	(18.2)	-217.9	(153.1)
4 crops	-27.9	(34.1)	2.4	(3.2)	4.2	(3.5)	0.9	(2.5)	-35.1	(34.3)
6 crops	36.1**	(15.9)	-0.3	(1.4)	-1.4	(2.4)	1.0	(1.6)	36.8**	(16.6)
7 crops or more	48.9**	(24.3)	-2.2	(1.9)	-5.3	(4.3)	-1.5	(2.8)	57.7**	(25.5)
Sample mean	2,424		294		161		96		1,874	

Note. Symbol "**", respectively "*", indicates that the parameter is significantly estimated at the 5%, respectively 10%, level. Estimated standard deviations of the parameter estimates are in parentheses.

As in the case of pre crop effects, crop acreage diversity effects are mostly due to effects on yield levels. Crop acreage diversity also significantly impacts herbicide uses from a statistical viewpoint although the related economic effects are limited. For instance, increasing the number of grown crops from 3 to 5 leads to savings in herbicide expenses on rapeseed of 11 €/ha on average, which equals to 11% of herbicide costs on rapeseed on average.

6. Concluding remarks

The main objective of this article is to estimate effects of crop diversity on yields and input uses. Because usually available datasets lack information regarding crop sequence acreages, we combine farm accounting data with IACS data, enriched with soil quality and weather data, and devise statistical models of yields and input uses. These models, which are defined as simultaneous equation systems, account for both input use endogeneity and unobserved heterogeneity of farms and farmers. In our application considering major arable crops in the Marne area, pre crops effects on yield levels are estimated relatively accurately and are generally consistent with the rankings provided by crop production experts. Estimated pre crop effects on input uses are small, suggesting that farmers tend to downplay them when deciding their chemical input use levels. Our results also show that crop acreage diversity, at least when described by a suitable set of indicators, increases yield levels and reduces pesticide uses, herbicide uses in particular. Taken together our results uncover statistically significant albeit economically limited effects of pre crops and crop acreage diversity on crop gross margins, at least in the economic context prevailing from 2008 to 2014.

Crop sequence acreages in our dataset are highly concentrated on the most profitable ones. This demonstrates that farmers' crop sequence acreages are economically rational but this also underlies an important drawback of farm data for estimating pre crop effects since it implies that pre crop effects can only be estimated for a limited number of previous crops and only for major crops. We interpret the fact that estimated effects on input uses are small because of farmers' neglecting these effects. Yet, other explanations can be put forward, including measurement errors in crop sequence acreages.

Although our application reveals statistically significant crop acreage diversity effects on both crop yield and herbicide use levels, our approach does not enable us to disentangle (long run) cropping system effects from (current) spatial crop diversity effects. Our constructing the crop acreage diversity indicators based on lagged acreage and results obtained by using the farm crop acreage histories suggest, but do not prove, that our estimated crop acreage diversity effects mostly capture cropping system effects. More generally, our approach identifies crop diversity effects by comparing the production choices and performances of farms characterized by heterogeneous crop acreages. Although we control for confounding factors by relying on detailed soil property measures, we cannot control for farmer skill heterogeneity that may impact both crop level production choices and crop acreage choices.

Our empirical results tend to show that crop diversification positively impacts the yield levels of major crops, that is to say of the crops that contribute the most to farmers' revenue– *i.e.*, straw cereals, rapeseed and sugar beet in our application. They also suggest that crop diversification has limited effects on chemical input uses on these major crops. Despite their being significant from a statistical viewpoint, the agronomic effects revealed by our study have relatively small impacts on the returns of major crops. This may explain why EU farmers keep on using relatively specialized crop rotations and crop acreages. Typical diversification crops such as grain legumes are minor crops in the UE, mostly due to their insufficient profitability in comparison to that of major crops (*e.g.*, Bues *et al*, 2013; Zander *et al*, 2016). The positive effects of crop diversity on the economic returns of major arable crops are unlikely to suffice for covering the opportunity cost of inserting typical diversification crops in otherwise specialized crop mixes (*e.g.*, Carpentier and Sodjahin, 2021).

Given our results, policy measures aimed to foster crop diversification are unlikely to significantly reduce chemical input uses on major crops if they are not supplemented by measures specifically aimed to reduce the uses of these inputs. Significant increases in the prices of pesticides and chemical fertilizers may partly solve this issue. First, this would increase the value of the chemical input use reductions on the major crops permitted by crop diversification and, as a result, could lead farmers to pay more attention to these reductions.

Second, this would reduce the impact of the opportunity cost of inserting legumes in farmers' crop mixes since these typical diversification crops do not require nitrogen fertilization. Setting incentive taxes on chemical inputs would, however, significantly impact arable crop producers' income and calls for specific measures aimed to neutralize the income effects of the considered taxing scheme.

Despite the limitations of our modelling framework and of the information content of our dataset, we are confident in the internal validity of our empirical measures of the pre crop and crop acreage diversity effects on yield and chemical input use levels in the arable crop sector of the Marne area. The external validity of our results is, however, more debatable since our case study displays salient specific features. The Marne area counts among the most productive arable crop production basin in the EU and farmers in this area rely on both relatively high yielding cropping practices and relatively diversified cropping systems. Yet pre crop effects, break crop ones in particular, may be less pronounced in diversified crop rotations than they are in specialized ones. Of course, further investigations, by economists and agronomists, are required for supporting or refuting these hypotheses.

References

Ackerberg, D.A., Caves, K. and Frazer, G. (2015). Identification Properties of Recent Production Function Estimators. *Econometrica* 83: 2411-2451.

Adeux, G., Munier-Jolain, N., Meunier, D., Farcy, P., Carlesi, S., Barberi, P. and Cordeau, S. (2019). Diversified grain-based cropping systems provide long-term weed control while limiting herbicide use and yield losses. *Agronomy for Sustainable Development* 39:42.

Arellano, M. and Bonhomme, S. (2011). Nonlinear Panel Data Analysis. *Annual Review of Economics* 35(1): 395-424.

Arrouays, D., Poggio, L., Salazar Guerrero, O. and Mulder, V.L. (2020). Digital Soil Mapping and GlobalSoilMap. Main advances and ways forward. *Geoderma Regional* 21:e000265.

Bareille, F. and Letort, E. (2018). How do farmers manage crop biodiversity? A dynamic acreage model with productive feedback. *European Review of Agricultural Economics* 45: 617-639.

Blundell, R. and Bond, S. (2000). GMM Estimation with persistent panel data: an application to production functions. *Econometric Reviews* 19: 321-340.

Bowles, T.M., Mooshammer, M., Socolar, Y., Calderon, F., Cavigelli, M.A., Culman, S.W., Deen, W., Drury, C.F., Garcia y Garcia, A., Gaudin, A.C.M., Harkcom, W.S., Lehman, R.M., Osborne, S.L., Robertson, G.P., Salerno, J., Schmer, M.R., Strock, J. and Grandy, A.S. (2020). Long-term evidence shows that crop rotation diversification increases agricultural resilience to adverse growing conditions in North America. *One Earth* 2: 284-293.

Bozzola, M. and Smale, M. (2020). The welfare effects of crop biodiversity as an adaptation to climate shocks in Kenya. *World Development* 135: 105065.

Breitung, J. and Salish, N. (2021). Estimation of heterogeneous panels with systematic slope variations. *Journal of Econometrics* 220: 399-415.

Bues, A., Preißel, S., Reckling, M., Zander, P., Kuhlman, T., Topp, K., Watson, C., Lindström, K., Stoddard, F.L. and Murphy-Bokern, D. (2013). *The environmental role of protein crops in the new common agricultural policy*. European Parliament No. 2012-067.

Carpentier, A. and Sodjahin, I.H.R. (2021). On the economics of crop rotation diversification: pre crop effects, cropping system effects and crop diversification opportunity costs. *Working Paper SMART, forthcoming*.

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

Chavas, J.-P. (2009). On the Productive Value of Biodiversity. *Environmental and Resource Economics* 42(1): 109-131.

Chavas, J.-P. and Di Falco, S. (2012). On the Role of Risk Versus Economies of Scope in Farm Diversification With an Application to Ethiopian Farms. *Journal of Agricultural Economics* 63: 25-55.

Chavas, J.-P. and Kim, K. (2007). Measurement and Sources of Economies of Scope. *Journal of Institutional and Theoretical Economics* 163: 411-427.

Chavas, J.-P. and Kim, K., (2010). Economies of diversification: A generalization and decomposition of economies of scope. *International Journal of Production Economics* 126: 229-235.

Chikowo, R., Faloya, V., Petit, S. and Munier-Jolain, N.M. (2009). Integrated Weed Management systems allow reduced reliance on herbicides and long-term weed control. *Agriculture, Ecosystems and Environment* 132: 237-242.

Colbach, N. and Cordeau, S. (2018). Reduced herbicide use does not increase crop yield loss if it is compensated by alternative preventive and curative measures. *European Journal of Agronomy* 94: 67-78.

Colbach, N., Petit, S., Chauvel, B., Deytieux, V., Lechenet, M., Munier-Jolain, N. and Cordeau, S. (2020). The pitfalls of relating weeds, herbicide use, and crop yield: Don't fall into the trap! A Critical Review. *Frontiers in Agronomy* 2: 615470.

Coulter, J.A., Sheaffer, C.C., Wyse, D.L., Haar, M.J., Porter, P.M., Quiring, S.R. and Klossner, L.D. (2011). Agronomic performances of cropping systems with contrasting crop rotations and external inputs. *Agronomy Journal* 103(1): 182-192.

Davis, A.S., Hill, J.D., Chase, C.A., Johanns, A.M. and Liebman, M. (2012). Increasing Cropping System Diversity Balances Productivity, Profitability and Environmental Health. *PLoS ONE* 7(10):e47149.

Delaune, T., Ouattara, M.S., Ballot, R., Sausse, C., Felix, I., Maupas, F., Chen, M., Morison, M., Makowski, D. and Barbu, C. (2021). Landscape drivers of pests and pathogens abundance in arable crops. *Ecography* 44: 1429-1442.

Delyon, B., Lavielle, M. and Moulines, E. (1999). Convergence of a Stochastic Approximation Version of the EM Algorithm. *The Annals of Statistics* 27(1): 94-128.

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

Di Falco, S. (2012). On the value of agricultural biodiversity. *Annual Review of Resource Economics* 4(1): 207-223.

Di Falco, S. and Chavas, J.-P. (2008). Rainfall Shocks, Resilience, and the Effects of Crop Biodiversity on Agroecosystem Productivity. *Land Economics* 84: 83-96.

Di Falco, S., Penov, I., Aleksiev, A. and van Rensburg, T.M. (2010). Agrobiodiversity, farm profits and land fragmentation: Evidence from Bulgaria. *Land Use Policy* 27: 763-771.

Di Falco, S. and Perrings, C. (2005). Crop biodiversity, risk management and the implications of agricultural assistance. *Ecological Economics* 55: 459-466.

Duru, M., Therond, O., Martin, G., Martin-Clouaire, R., Magne, M.-A, Justes, E., Journet, E.-P., Aubertot, J.-N., Savary, S., Bergez, J.-E. and Sarthou, J.-P. (2015). How to implement biodiversity-based agriculture to enhance ecosystem services: a review. *Agronomy for Sustainable Development* 35(4):1259-1281.

European Commission (2017). *CAP Explained. Direct payments for farmers 2015-2020.* Directorate-General for Agriculture and Rural Development (European Commission). https://doi.org/10.2762/149509

European Commission (2019). Integrated Administration and Control System (IACS) [WWW Document]. URL https://ec.europa.eu/info/food-farming-fisheries/key-policies/common-agricultural-policy/income-support/controls-and-transparency/managing-payments_en (accessed on November 6, 2019).

Feng, H., Wang, T., Osborne, S.L. and Kumar, S. (2021). Yield and economic performance of crop rotation systems in South Dakota. *Agrosystems, Geosciences and Environment* 4:e20196.

Frisvold, G. B. (2019). How low can you go? Estimating impacts of reduced pesticide use. *Pest management science*, *75*(5):1223-1233.

Griliches, Z. and Mairesse, J. (1995). Production functions: the search for identification. NBER Working Paper N° 5067.

Groom, B. and Pereira Fontes, F. (2021). Revisiting the link between cereal diversity and production in Ethiopia. *Q Open* 1: 1-21.

Guyomard, H., Bureau, J.-C. Chatellier, V., Detang-Dessendre, C., Dupraz, P., Jacquet, F., Reboud, X., Requillart, V., Soler, L.-G. and Tsyebaert, M. (2020). *Research for AGRI Committee - The Green Deal and the CAP: policy implications to adapt farming practices and to preserve the EU's natural resources*. European Parliament, Policy Department for Structural and Cohesion Policies, Brussels.

Hufnagel, J., Reckling, M. and Ewert, F. (2020). Diverse Approaches to Crop Diversification in Agricultural Research. A Review. *Agronomy for Sustainable Development* 40(14).

Hunt, N.D., Hill, J.D. and Liebman, M. (2017). Reducing freshwater toxicity while maintaining weed control, profits, and productivity: effects of increased crop rotation diversity and reduced herbicide usage. *Environmental Science and Technology* 51:1707-1717.

Ikerd, J.E. (1993). The need for a systems approach to sustainable agriculture. *Agriculture, Ecosystems and Environment* 46:147-160.

Jeuffroy, M.-H., Biarnès, V., Cohan, J.-P., Corre-Hellou, G., Gastal, F., Jouffret, P., Justes, E., Landé, N., Louarn, G., Plantureux, S., Schneider, A., Thiébeau, P., Valantin-Morison, M. and Vertès, F. (2015). Chapitre 3. Performances agronomiques et gestion des légumineuses dans les systèmes de productions végétales. In Schneider, A. and Huyghes, C. (eds), *Les légumineuses pour des systèmes agricoles et alimentaires durables*: 139-224.

Just, R.E. and Pope, R.D. (2001). Chapter 12 The agricultural producer: Theory and statistical measurement, in: *Handbook of Agricultural Economics, Agricultural Production.* Elsevier, 629-741.

Khakbazan, M., Gan, Y., Bandara, M. and Huang, J. (2020). Economics of pulse crop frequency and sequence in a wheat-based rotation. *Agronomy Journal* 112: 2058-2080.

Koutchadé, O.P., Carpentier, A. and Femenia, F. (2018). Modeling Heterogeneous Farm

Responses to European Union Biofuel Support with a Random Parameter Multicrop Model. *American Journal of Agricultural Economics* 100(2): 434-455.

Koutchadé, O.P., Carpentier, A. and Femenia, F. (2021). Modeling Corners, Kinks, and Jumps in Crop Acreage Choices: Impacts of the EU Support to Protein Crops. *American Journal of Agricultural Economics* 103(4): 1502-1524.

Kremen, C. and Miles, A. (2012). Ecosystem Services in Biologically Diversified versus Conventional Farming Systems: Benefits, Externalities, and Trade-Offs. *Ecology and Society* 17(4):40.

Kuhn, E. and Lavielle, M. (2005). Maximum likelihood estimation in nonlinear mixed effects models. *Computational Statistics and Data Analysis* 49: 1020-1038. https://doi.org/10.1016/j.csda.2004.07.002

Lavielle, M. (2014). *Mixed Effects Models for the Population Approach: Models, Tasks, Methods and Tools*. CRC Press.

Lin, B.B. (2011). Resilience in Agriculture through Crop Diversification: Adaptive Management for Environmental Change. *BioScience* 61: 183-193.

Liu, K., Johnson, E.N., Blackshaw, R.E., Hossain, Z. and Gan, Y. (2019). Improving the productivity and stability of oilseed cropping systems through crop diversification. *Field Crops Research* 237: 65-73.

Louhichi, K., Ciaian, P., Espinosa, M., Perni, A. and Gomez y Paloma, S. (2018). Economic impacts of CAP greening: application of an EU-wide individual farm model for CAP analysis (IFM-CAP). *European Review of Agricultural Economics* 45(2): 205-238.

Loyce, C.H. and Meynard, J.-M. (1997). Low input wheat management techniques are more efficient in ethanol production. *Industrial Crops and Products* 6(3-4): 271-283.

Loyce, C.H., Meynard, J.-M., Bouchard, C., Rolland, B., Lonnet, P., Bataillon, P., Bernicot, M.-H., Bonnefoy, M., Charrier, X. and Debote, B. (2008). Interaction between cultivar and crop management effects on winter wheat diseases, lodging, and yield. *Crop protection* 27(7): 1131-1142.

Loyce, C.H., Meynard, J.-M., Bouchard, C., Rolland, B., Lonnet, P., Bataillon, P., Bernicot, M.-H., Bonnefoy, M., Charrier, X. and Debote, B. (2012). Growing winter wheat cultivars under different management intensities in France: A multicriteria assessment based on economic, energetic and environmental indicators. *Field crops research* 125: 167-178.

Magrini, M.-G., Anton, M., Cholez, C., Corre-Hellou, G., Duc, G., Jeuffroy, M.-H., Meynard, J.-M., Pelzer, E., Voisin, A.-S. and Walrand, S. (2016). Why are grain-legumes rarely present in cropping systems despite their environmental and nutritional benefits? Analyzing lock-in in the French agrifood system. *Ecological Economics* 126: 152-162.

Martin, P., Levavasseur, F., Bouty, C. and Scheurer, O. (2017). RPG Explorer Version 1.8.148 Notice d'utilisation.

McLachlan, G. J. and Krishnan, T. (2007). *The EM Algorithm and Extensions*. John Wiley & Sons.

Meynard, J.M., Messéan, A., Charlier, A., Charrier, F., Farès, M., Le Bail, M. and Savini, I. (2013). *Crop Diversification: Obstacles and Levers*. INRA: Paris, France.

Morel, K., Revoyron, E., San Cristobal, M. and Baret, P. (2020). Innovating within or outside dominant food systems? Different challenges for contrasting crop diversification strategies in Europe. *PLoS ONE* 15(3), 24p.

Mortensen, D.A. and Smith, R.G. (2020). Confronting Barriers to Cropping System Diversification. *Frontiers in Sustainable Food Systems* 4:564197.

Mundlak, Y. (1996). Production Function Estimation: Reviving the Primal. *Econometrica* 64(2): 431-438.

Mundlak, Y. (2001). Chapter 1 Production and supply, in: *Handbook of Agricultural Economics, Agricultural Production*. Elsevier, 3-85.

Nave, S., Jacquet, F. and Jeuffroy, M.-H. (2013). Why wheat farmers could reduce chemical inputs: evidence from social, economic, and agronomic analysis. *Agronomy for Sustainable Development* 33: 795-807.

Omer, A., Pascual, U. and Russell, N.P. (2007). Biodiversity conservation and productivity in intensive agricultural systems. *Journal of Agricultural Economics* 58(2): 308-329.

Preißel, S., Reckling, M., Schläfke, N. and Zander, P. (2015). Magnitude and farm-economic value of grain legume pre-crop benefits in Europe: a review. *Field Crops Research* 175:64-79.

Sharma, G., Shrestha, S., Kunwar, S. and Tseng, T.-M. (2021). Crop diversification for improved weed management: A review. *Agriculture* 11: 461.

Smith, R.G. and Gross, K.L. (2007). Assembly of Weed Communities along a Crop Diversity Gradient. *Journal of Applied Ecology* 44: 1046-1056.

Smith, E.G., Zentner, R.P., Campbell, C.A., Lemke, R. and Brandt, K. (2017). Long-Term Crop Rotation Effects on Production, Grain Quality, Profitability, and Risk in the Northern Great Plains. *Agronomy Journal* 109(3): 957-967.

Suri, T. (2011). Selection and comparative advantage in technology adoption. *Econometrica* 79(1): 159-209.

Tembo, G., Brorsen, B. W., Epplin, F. M., and Tostão, E. (2008). Crop input response functions with stochastic plateaus. *American Journal of Agricultural Economic* 90(2): 424-434.

Thérond, O., Duru, M., Roger-Estrade, J. and Richard, G. (2017). A new analytical framework of farming system and agriculture model diversities. A review. *Agronomy for Sustainable Development* 37(21).

Thiessen Martens, J.R., Entz, M.H. and Wonneck, M.D. (2015). Review: Redesigning Canadian prairie cropping systems for profitability, sustainability, and resilience. *Canadian Journal of Plant Science* 95:1049-1072.

Watson, C.A., Reckling, M., Preißel, S., Bachinger, J., Bergkvist, G., Kuhlman, T., Lindström, K., Nemecek, T., Topp, C.F.E., Vanhatalo, A., Zander, P., Murphy-Bokern, D. and Stoddard, F.L. (2017). Grain Legume Production and Use in European Agricultural Systems. *Advances in Agronomy* 144(1):235-303.

Wooldridge, J.M. (2005a). Fixed effects and related estimators in correlated random coefficient and treatment effect panel data models. *Review of Economics and Statistics* 87: 385-390.

Wooldridge, J.M. (2005b). Unobserved heterogeneity and estimation of average partial effects. In: Andrews, D.W.K., Stock, J.H. (Eds.), *Identification and Inference for Econometric Models*: A Festschrift in Honor of Thomas J. Rothenberg. Cambridge University Press, Cambridge, 27-55.

Wooldridge, J.M. (2010). *Econometric Analysis of Cross Section and Panel Data*. MIT Press. 2nd edition.

Wooldridge, J.M. (2019). Correlated random effects models with unbalanced panels. *Journal of Econometrics* 211: 137-150.

Wu, C.F.J. (1983). On the Convergence Properties of the EM Algorithm. *The Annals of Statistics* 11(1): 95-103.

Zander, P., Amjath-Babu, T.S., Preißel, S., Reckling, M., Bues, A., Schläfke, N., Kuhlman, T., Bachinger, J., Uthes, S., Stoddard, F., Murphy-Bokern, D. and Watson, C. (2016). Grain legume decline and potential recovery in European agriculture: a review. *Agronomy for Sustainable Development* 36(2):26-36.

Zellner, A. (1962). An Efficient Method of Estimating Seemingly Unrelated Regressions and Tests for Aggregation Bias. *Journal of the American Statistical Association* 57: 348-368.

Zentner, R.P., Wall, D.D., Nagy, C.N., Smith, E.G., Young, D.L., Miller, P.R., Campbell, C.A., McConkey, B.G., Brandt, S.A., Lafond, G.P., Johnston, A.M. and Derksen, D.A. (2002). Economics of crop diversification and soil tillage opportunities in the Canadian prairies. *Agronomy Journal* 94(2): 216-230.