



# Unintended Consequences of Environmental Policies: the Case of Set-aside and Agricultural Intensification

Raja Chakir<sup>1</sup> · Alban Thomas<sup>2</sup>

Received: 19 October 2020 / Accepted: 20 December 2021 / Published online: 29 January 2022  
© The Author(s), under exclusive licence to Springer Nature Switzerland AG 2022

## Abstract

The objective of this paper is to evaluate the indirect impacts of a set-aside policy on the environment through crop intensification. We estimate a structural, multi-output model on a panel of French farmers from 2006 to 2010, accounting for multivariate selection (corner solutions) on crops and land use. We use a parametric and a semi-nonparametric (based on sieve estimation) version of a quasi-maximum likelihood (QML) estimator, compare their goodness of fit and parameter efficiency and check for closeness with a test for non-nested models. We then perform a simulation experiment based on elasticity estimates for set-aside area and input demand, to evaluate intensive margin effects of the set-aside policy. Results show that a set-aside subsidy can provide farmers with incentives to intensify their production, leading to potential adverse environmental effects.

**Keywords** Set-aside · Land use · Fertilizer and pesticide input demand · Corner solutions · Semi-nonparametric estimation

**JEL Classification** Q12 · C33 · C34

## 1 Introduction

Land set-aside (fallow) schemes, involving farmers being paid to take land out of production, are widely used in the European Union (EU) and the USA as an agricultural policy tool. In many cases, the initial aim of this measure was to reduce excess supply of cereals while lowering the level of public agricultural stocks. It was later complemented by tools to promote the development of non-food crops and maintain good environmental status. In the USA, the Conservation Reserve Program (CRP) was introduced in 1985 as a voluntary set-aside program designed to control crop overproduction, reduce soil erosion, improve water quality, and provide wildlife habitat by taking vulnerable agricultural land out of production. In the EU, compulsory set-aside was

one of the most important measures introduced at the time of the 1992 Common Agricultural Policy (CAP) reform. In 2008, the policy package associated with the CAP Health Check abolished set-aside for arable crops, but farmers could continue to set land aside on a voluntary basis while adopting agri-environmental schemes with cross-compliance. With the reform of the CAP for 2014–2020, the European Commission introduced the Green Direct Payment program that links direct payments to farmers to requirements for mandatory “greening” farming practices. These “Greening” practices include: (1) crop diversification, (2) the maintaining of permanent grassland and (3) Ecological Focus Areas (EFAs). EFAs are areas considered to have environmental benefits such as fallow land, margins catch crops, green cover and nitrogen-fixing crops. From 2015 on, farmers with more than 15 ha must have at least 5% of their land as an EFA. This “greening” of the CAP was described as a move back to compulsory set-aside<sup>1</sup>.

The objective of this paper is to evaluate the potential impacts of a set-aside policy on the environment, more precisely in terms of intensification of input use (fertilizer and pesticide) on cultivated crop area (intensive margin).

✉ Raja Chakir  
Raja.Chakir@inrae.fr

Alban Thomas  
Alban.Thomas@inrae.fr

<sup>1</sup> Université Paris-Saclay, INRAE, AgroParisTech, Paris-Saclay Applied Economics, F-91120 Palaiseau, France

<sup>2</sup> INRAE, Toulouse School of Economics-Research, University of Toulouse, 31000 Toulouse, France

<sup>1</sup> See: <http://capreform.eu/assessment-of-the-commission's-proposal-for-an-obligatory-set-aside-programme/>

To capture these effects, we need to account for changes in the distribution of crops, because fertilizer and pesticide requirements are often heterogeneous across crops, as are their potential environmental impacts. This means that using individual farm data and for each crop is preferable in terms of policy impact evaluation. However, this type of data raises corner solution issues that are critical in empirical work at the individual farm level.

We apply a dedicated estimation procedure to estimate acreage and agricultural practice responses to the set-aside policy, for a sample of farmers from the *département* of Meuse (eastern France), observed from 2006 to 2010. We estimate a multi-output profit function based on a panel of individual farmers, controlling for multiple selection using both a parametric and a semi-nonparametric quasi-maximum likelihood (QML) estimator. The purpose of this econometric strategy is to check for possible deviations from heteroskedasticity of the parametric QML estimator, by comparing it with a more flexible semi-nonparametric estimator. The latter is based on sieve estimation [1] that allows for consistent, and in general more efficient, estimates than semiparametric estimation based on, e.g., kernel nonparametric estimation of the underlying distribution of error terms. Parameters estimated from a structural multicrop system of equations allow us to estimate the change in pesticide and fertilizer demand corresponding to a set-aside policy. We propose two indicators to measure these impacts: chemical input (pesticide and fertilizer) demand and intensity elasticities with respect to the set-aside subsidy rate.

**Paper Contributions** The present paper makes two contributions to the literature. The first contribution is methodological and concerns the econometric strategy for dealing with multiple selection in a micro-panel sample of farms. Our contribution in this regard is to propose a consistent and flexible estimator for output and input decisions from individual farm data. To allow for correlation between (possibly censored) equations in the multivariate Tobit framework while avoiding multiple integration, we consider a QML approach, based on pairwise joint probabilities of structural equations. Yen et al. [2] and Fezzi and Bateman [3] use a (parametric) QML estimator, which is known to be consistent under the assumption that the conditional mean of the model is correctly specified and that the density used for QML belongs to the linear exponential family [4]. However, the QML may not be consistent if higher moments of the error distribution are introduced in the QML criterion but are misspecified (typically, variance–covariance terms). To address this issue, we propose in this paper a semi-nonparametric version of the QML estimator applied to the multivariate selection model, which allows us to relax both distributional and homoscedasticity assumptions. By doing so, we go beyond the

procedure proposed by [5], which is consistent only if the selection equations are uncorrelated. Our estimator is based on semi-nonparametric sieve estimation, which is faster to compute than semiparametric estimators based on kernel approximations (as in, e.g., [6]). To the best of our knowledge, such semi-nonparametric methods to obtain output supply and input demand estimates from farm-level data have not yet been used in the literature. Because crops are different in terms of input requirements and environmental impacts, accounting for such heterogeneity at the farm level is likely to produce more precise estimates of farm input demands. This will allow us to more precisely estimate the environmental impacts of the set-aside policy in terms of agrochemical input intensification.

The second contribution of our paper concerns the *ex ante* environmental assessment of a land set-aside policy, in terms of agro-chemical input intensification. We contribute to the debate on the environmental impacts of greening policies in general, and on set-aside policies in particular. Indeed, some studies show that set-aside land supports biodiversity with much higher population densities and numbers of species of birds, insects, spiders, and plants [7]. This contradicts studies in which set-aside land has been found to be ineffective and inefficient for conservation [8]. More recently, the report of the [9] concludes that “*greening, as currently implemented, is unlikely to significantly enhance the CAP’s environmental and climate performance*”. This is mainly due, according to the same report, to the fact that “*Greening lacks a fully developed intervention logic with clearly defined, ambitious targets...*”. Indeed, farmers were not given any environmental target to achieve through the greening reform.

Our paper proposes a quantitative assessment of a “greening” policy, which aims to increase the set-aside area by 5%. The first feature of this contribution consists of evaluating the indirect impact of this set-aside policy on crop intensification using two indicators: chemical input demand elasticities and a new indicator, namely input demand intensity elasticities, both with respect to the set-aside subsidy rate. Such a complementary indicator is better suited, in our opinion, to agricultural settings where input use per unit of land is more relevant than total input demand. The second feature of this contribution is to simulate the impacts, in terms of fertilizer and pesticide demand variation, of a policy aimed at increasing set-aside area by 5 percent, and to compute the level of the input tax that will be necessary to cope with the increased demand. Reducing fertilizer and pesticide use is an environmental objective that has been explicitly introduced in the recent *Farm to Fork Strategy* of the European Commission, as part of the European Green Deal initiative. This strategy aims to reduce the use of fertilizers and pesticides in EU agriculture by at least 20% and 50%, respectively, by 2030. It is important to note that the impact evaluation

**Table 1** Past and ongoing Common Agricultural Policy (CAP) measures regarding set-aside

Year - Period	Set-aside reform	Main objectives
1988–1989	Voluntary set-aside	Prevent overproduction of some commodities
1992 MacSharry reform	Compulsory set-aside	Limit stocking levels
Agenda 2000	Decoupling	
2003–2006	Attribution of set-aside payment on an historical basis	Limit supply and provide better environmental efficiency
2008 CAP health check	Abolition of the set-aside obligation	
2013	Green direct payments and Ecological Focus Area (EFA)	Reduce biodiversity loss and GHG emissions
2023–2027	At least 10% of agricultural areas devoted to non-productive features	Improve farmland biodiversity

concerns only potential benefits (or negative externalities) from a reduction in nutrient load, following a change in land use and input use (fertilizer, pesticide). Strictly speaking, we are not evaluating actual environmental effects of the set-aside policy on, e.g., water quality, greenhouse gas emissions (nitrous oxide from chemical fertilizer) and biodiversity, but rather the potential of such policy to limit chemical input use at the farm level.

The paper is organized as follows: Section 2 discusses land set-aside policies in the context of the European CAP and presents a literature survey of both the impacts of set-aside policies and econometric models with corner solutions. Section 3 presents a production model based on a multicrop framework and econometric methods to deal with corner solutions. Parametric and semi-nonparametric estimators are discussed in the context of panel data. Section 4 describes data used in our econometric analysis and discusses estimation results. In Sect. 5, we discuss results of chemical input demand and intensity elasticities and simulate the impact of a policy consisting of a 5-percent set-aside rate as an agricultural greening measure and calculate the input tax on fertilizer and pesticide that would be necessary to cope with increased demand for these inputs. Section 6 concludes the paper.

## 2 Policy Context and Related Literature

### 2.1 CAP and Set-aside Policy Context

Compulsory set-aside was one of the most important measures introduced in the European Union (EU) at the time of the 1992 reform of the Common Agricultural Policy (CAP), which introduced a new support system for producers of cereals, oilseed and protein crops. See Table 1 for a summary of CAP measures regarding set-aside policies. Farmers with production greater than 92 tons were eligible for set-aside payments. In order to alleviate their revenue decrease due to the compulsory set-aside, farmers were allowed to cultivate energy crops (diester from rapeseed in our data)

on set-aside land without losing the subsidy [10]. Among the many changes to set-aside rules during the period 1993–2007, the major ones concerned adjustments to the rate of compulsory set-aside, the introduction of voluntary set-aside against payment, and the possibility of a fixed instead of a rotational set-aside, which was the only form available at the outset. In 2008, the policy package associated with the Health Check of the CAP abolished both the energy crop scheme and the compulsory set-aside scheme. However, farmers could continue to set land aside on a voluntary basis, while adopting agri-environmental schemes with cross-compliance. The eligibility requirement for payment in that case is that at least 5 percent of land should be within an ecological focus area. Moreover, eligibility conditions remain the same regarding crop area: farms are eligible for area-based payment only if the area planted with cereals, oilseed, and protein crops is greater than 0.3 hectares, under European directive CE/1973/2004 (October 29, 2004). Since 2008, the ceiling for the maximum voluntary set-aside area eligible for CAP payment is 1/9<sup>th</sup> of the arable area. It can be extended to 25 percent for arable crops (cereals, oilseed, and protein crops) for energy, chemical use, or animal feed. In practice, for rapeseed production, farmers were eligible for area-based payment under CAP after the 1992 reform, while voluntary set-aside over and above the compulsory rate was allowed. After 2008, rapeseed for diester production was also possible on set-aside areas, with the same subsidy rate as for non-set-aside areas, i.e., a different subsidy rate from the “agronomic” set-aside rate. In this case, farmers have to show evidence of a farming contract with energy or industrial buyers.

In December 2013, the EU enacted the CAP for the period 2014–2020 where, in addition to the Basic Payment Scheme, each farmer now receives a Green Direct Payment per hectare for respecting specific greening agricultural practices to reduce biodiversity loss and greenhouse gas emissions [11]. Member States are requested to use 30 percent of their national budget in order to contribute to this program. The three greening measures are as follows: (i) maintaining permanent grassland; (ii) crop diversification;

and (iii) maintaining an Ecological Focus Area (EFA) of at least 5 percent of the arable area of the holding for farms with an area larger than 15 hectares (excluding permanent grassland), i.e., field margins, hedges, trees, fallow land, landscape features, biotopes, buffer strips, afforested area. Indeed, set-aside originally introduced for supply control purposes could have important environmental benefits, especially where land was left fallow [12]. However, the literature review by [13] shows that the CAP greening measures could have mixed effects on ecosystem services, as they can lead to intensification on cultivated land.

In May 2020, the European Commission officially presented a “Farm to Fork” strategy as part of the European Green Deal initiative. This strategy aims to make “the entire food chain from production to consumption more sustainable and neutral in its impact on the environment”. Among the targets considered to reach these objectives, the strategy mentions a 50% reduction in the use of chemical pesticide and a reduction in fertilizer use of at least 20%, both by the year 2030.

For the last CAP reform which will be implemented from 2023, Member States are required to “contribute” to the Green Deal via national agricultural policies. For example, among Good Agricultural and Environmental Conditions (GAEC), GAEC9 stipulates that all farms should have at least 10% of their agricultural area devoted to non-productive features in order to improve farmland biodiversity [14].

## 2.2 Related Literature

### 2.2.1 Literature on the Impacts of Set-aside Policies

By modifying the opportunity cost of farm land, a set-aside policy is expected to modify farmers’ production decisions in terms of crop choice and input use. However, the impact of set-aside policies on input use and intensification remains unclear: on the one hand, removing a proportion of land from production might reduce input use and increase extensification [15, 16]. On the other hand, set-aside might have adverse effects on water and soil quality if input use (fertilizer, pesticide) increases in order to balance the reduction in cultivated land. The impact of a set-aside policy therefore depends on the degree of use of the intensive margin (i.e., intensification of agrochemical input on the same crop land) vs. the extensive margin (adapting the distribution of land to crops, some being more intensive in chemical inputs than others, but leaving input level per hectare unchanged) used by the farmer in his production decisions. See [16–22] for more details on slippage effects. Gorddard [23] explores the relationships between crop production, land allocation, and input-use decisions, exploring the consequences of assuming production of multiple crops

is non-joint but is subject to a constraint on total land area. His results illustrate that the effectiveness of environmental policies (aimed at modifying input use) may be affected by land allocation decisions. In the case of the CRP, early analysis found that contracts were targeted to reduce production, rather than achieving environmental benefits [24]. Hendricks and Er [25] show that the government has adjusted CRP acreage over time in response to changes in market conditions but not to environmental impacts<sup>2</sup>. In the case of the new EFA policy implemented in the EU and aimed at preserving biodiversity, the intensity of the impact depends on the type of EFA (fallow, grassland), the specific environmental issue considered (water pollution, soil pollution, biodiversity loss), and site-specific environmental conditions.

The impact of set-aside policies on the environment is not only associated with input intensification and its consequences on nutrient load, but also on climate change and biodiversity, although such environmental dimensions are more difficult to directly connect to changes in input use. For example, [26] present an impact evaluation of a set-aside policy in Finland, by combining a production model with area-based set-aside payments and estimates of monetary benefits of reduced nutrient pollution. A cost–benefit analysis is then conducted on the nutrient load reductions attributable to CAP agri-environmental payments in Finland between 1996 and 2005. In any case, policy impacts of set-aside are also relevant to the debate over land sparing versus land sharing (see, e.g., [27]).

Relying on the intensive margin in cropping systems is one of the unintended impacts of set-aside policy, as is the slippage effect that was observed in the CRP in the USA [22]. Both impacts can be explained by the increased output prices associated with reduced production on set-aside land. In particular, input decisions may be modified following a change in the set-aside subsidy rate, if a) it is more profitable to rely on the extensive margin; and/or b) with the same distribution of crops, it pays to intensify production on a smaller proportion of land, using the intensive margin by increasing application rates of fertilizer and pesticide by the same proportion on all crops. These two changes can be measured by a combination of elasticities: agrochemical input demand and intensity of agrochemical input use, both with respect to the set-aside subsidy rate.

<sup>2</sup> The set-aside policy implemented in the EU is a little complex as farmers are allowed to plant bioenergy crops. In the USA, set-aside land, generally under the CRP program, is only allowed to grow trees or perennial grasses see [84]

## 2.2.2 Literature on Econometric Modelling of Corner Solutions

Although there have been major advances in the estimation of production technology, there remain some challenging econometric issues. One problem specifically relevant to agricultural land-use allocation in empirical studies is corner solutions. Corner solutions arise when it is optimal for the farmer not to grow a crop (or combination of crops). Using aggregate data on land allocation, for example at the regional level, we observe positive values for all (region-specific) crops, but this does not imply that all farmers grow all crops. When using individual farm-level data, we need to account for the fact that some farmers may choose not to grow some crops, thus possibly causing selection bias in the parameter estimates. For this reason, models that analyze the effects of agricultural policies should adopt an explicit methodology that accommodates and explains the existence of corner solutions in the context of micro-level farm data.

Most of the literature on multi-crop estimations relies on aggregate data or does not deal explicitly with corner solutions that occur in land-use decisions. Guyomard et al. [28] estimate a quadratic profit function with several crop groups and inputs using French aggregate data but do not discuss the issue of corner solutions in production. Moro and Sckokai [29] employ a normalized quadratic multi-output profit function on Italian FADN (Farm Accounting Data Network) data but do not exploit the panel data structure of their (individual) data set. Moreover, although they recognize the presence of possible sample selection when dealing with multiple crop groups, they do not control for these sample selection effects explicitly or consistently.

The most common technique used to estimate a structural model subject to censored observations is Tobit estimation. This was proposed in the econometric literature by [30] and has been widely utilized in the empirical literature on demand estimations. Although the Tobit model is useful, it is an *ad hoc* modification of the regression model, allowing it to be used in cases where there are observations “piled up” at some limiting value (usually zero), and it has no convincing behavioral theory foundation [31]. The standard solution to the problem of a censored dependent variable is to estimate a Tobit model using maximum likelihood (ML) or the [32] two-step method.

The pioneering works of [33] and [34] offer an economic interpretation of corner solutions and a direct and appropriate method for specifying the econometric model. They explain that the set of producer choices can be analyzed employing the Kuhn–Tucker conditions associated with the cost minimization program under the usual technical constraints and nonnegativity constraints on input demand. The implied fully structural approach in [35] and [34] to estimate

demand and take account of corner solutions is nonlinear simultaneous zero-censored equation models. When the number of equations is large, the subset of decision outcomes in the system likely to occur at kink points increases, requiring multiple integration for ML estimation.

In response to the issue of dimensionality when estimating demand systems with binding nonnegativity constraints, many strategies are adopted in the literature. Alternative estimation methods to the ML procedure include the maximum entropy estimator [36, 37], the two-step Tobit system [38], generalized method of moments (GMM) techniques [39]. Yen et al. [2] propose a QML approach which they claim is more efficient for small- to medium-sized samples. Comparison with other estimation methods shows that the QML and SML (simulated maximum likelihood) procedures produce remarkably similar demand parameter and elasticity estimates, whereas the results of [40]’s two-step estimator differ widely. Yen and Lin [41] propose a sample-selection alternative with more flexible parameterization (than the Tobit system), using the [40] two-step estimator. This approach is used by [5] in the multiple selection case. However, the sample selection system presents more prominent computational burdens than the Tobit system since the sample likelihood function contains probability integrals with dimensions as large as the number of selection equations for all sample observations.

Simulation-based estimation methods have been suggested to overcome this problem of high-dimensional numerical integration in multivariate limited dependent variable systems. These methods include simulated moments, simulated maximum likelihood, and simulated scores. The simulated ML approach is applied by [42] and [43]. Another way to overcome the computational issue is to use Bayesian methods. Millimet and Tchernis [44] use the Gibbs sampling technique with the data augmentation algorithm to solve both the dimensionality and coherency problems<sup>3</sup> while [45] employ the two-step approach proposed by [40] to take account of both censoring and unbalanced panel data structure. Platoni et al. [46] account also for the heteroscedastic structure of the error terms in the (second-step) estimation of the expectation-conditional maximization (ECM) model.

<sup>3</sup> Coherency conditions in a structural econometric modeling framework are conditions that ensure that the model is correctly specified from the theoretical point of view. For example, [85] show that a negative semi-definite parameter matrix is necessary and sufficient for the translog cost function to be concave. In most econometric models, estimating a cost function imposing the semi-negativity of the parameter matrix ensures that the coherency conditions to be respected.

### 3 The Model and Estimation Issues

#### 3.1 The Model

Consider a risk-neutral<sup>4</sup> farmer using  $K$  variable inputs  $x$  and a fixed but assignable factor (land) to produce  $C$  different crops, where  $c$  is the crop index,  $c = 1, \dots, C$ ,  $p_c$  is the price of crop  $c$ ;  $y_c$  is the output level of crop  $c$ ,  $w_k$  is the price of input  $k$ ;  $l_c$  is the land allocated to crop  $c$  and  $L$  is the total available land ( $\sum_{c=1}^C l_c = L$ ).  $\tau_c$  is the area-based (per hectare) subsidy rate for crop  $c$ . To simplify notation, we consider set-aside as a particular land use. Therefore, it is also indexed by  $c$ , although it is not a crop, technically speaking.

Maximizing profit over land and input choices, given the predetermined system of prices, subsidies, and total available land, will produce a unique set of solutions, provided that regularity conditions are satisfied for the profit function<sup>5</sup>. Following [47] and [28], the multi-crop profit function for a joint input technology given the fixed factor allocation (land) may be written as:

$$\Pi(p, w, \tau, L) = \max_{y,x,l} \left\{ \sum_{c=1}^C p_c y_c - \sum_{k=1}^K w_k x_k + \sum_{c=1}^C \tau_c l_c; \sum_{c=1}^C l_c = L; y \leq F(x, L) \right\}, \tag{1}$$

where  $y, x, w$ , and  $\tau$  are  $C$ -dimension vectors and  $y \leq F(x, L)$  represents the (multi-output) technological feasibility set.

Under partial decoupling of public payments to agriculture, the equation above has to be modified, to accommodate the situation where only a proportion of area-based payments is crop-specific. Let  $\phi, \phi \in [0, 1]$ , denote the proportion of area-based subsidies that is decoupled from production, so that  $100(1 - \phi)$  percent are still ‘‘coupled’’ and depend on the area allocated to production for each crop  $c$ .

Land set-aside as a particular land use can be divided into a compulsory part (set-aside obligation) and a voluntary component, defined as the fraction of land set-aside above

<sup>4</sup> Risk preferences are relevant to a discussion of agricultural production, but we do not consider an extension of our framework to accommodate farmer risk attitudes. A dedicated structural approach would be necessary, but also particularly difficult to adapt to multi-output production with selection, see [86] for an example. Moreover, the literature distinguishes two approaches relating to the modeling of pesticides: as ‘‘damage abatement inputs’’ or as ‘‘productive inputs’’. In this paper, we consider the second approach. Indeed, [87] compared these two approaches and showed that our approach, in which all inputs are treated as productive inputs, does not lead to larger estimates of the value of the marginal product of pesticides as it was hypothesized in the literature.

<sup>5</sup> We assume in this paper that farmers are rational profit-maximizing individuals. However, [88] discussed the fact that rationality assumptions can be challenged in the context of farmers’ decisions to adopt more sustainable practices (as in pesticide use decisions).

the minimum requirement of an agricultural policy. In our case, the area-based payment for set-aside is the same in both cases; we denote by  $G$  the land set-aside obligation, with unit payment  $g$  (same notation as in [28]). As a consequence, land used for crops or voluntary set-aside is a fraction of total farm land and is denoted  $L^* = L - G$ .

Farmer profit therefore includes a fixed payment denotes  $FP$ , consisting of area-based subsidies not coupled with production nor land use (proportional to  $\phi$ ) and the compulsory set-aside payment,  $gG$ .

Farmer’s profit then becomes

$$\begin{aligned} \Pi(p, w, \tau, L^*) &= \max_{y,x,l} \left\{ \sum_{c=1}^C p_c y_c - \sum_{k=1}^K w_k x_k \right. \\ &\quad \left. + (1 - \phi) * \sum_{c=1}^C \tau_c l_c + FP; \sum_{c=1}^C l_c \right. \\ &\quad \left. = L^*; y \leq F(x, L^*) \right\}. \end{aligned} \tag{2}$$

Following [5], the normalized quadratic profit function is written as:

$$\begin{aligned} \bar{\Pi} &= \alpha_0 + \sum_{c=1}^C \alpha_c \bar{p}_c + \sum_{k=1}^{K-1} \beta_k \bar{w}_k + \sum_{c=1}^C \gamma_c \bar{\tau}_c + \frac{1}{2} \sum_{c=1}^C \sum_{c'=1}^C \alpha_{cc'} \bar{p}_c \bar{p}_{c'} \\ &\quad + \frac{1}{2} \sum_{k=1}^{K-1} \sum_{k'=1}^{K-1} \beta_{kk'} \bar{w}_k \bar{w}_{k'} + \frac{1}{2} \sum_{c=1}^{C-1} \sum_{c'=1}^C \gamma_{cc'} \bar{\tau}_c \bar{\tau}_{c'} \\ &\quad + \sum_{k=1}^{K-1} \sum_{c=1}^C \delta_{ck}^{pw} \bar{p}_c \bar{w}_k + \sum_{c=1}^C \sum_{c'=1}^C \delta_{cc'}^{p\tau} \bar{p}_c \bar{\tau}_{c'} + \sum_{k=1}^{K-1} \sum_{c=1}^C \delta_{ck}^{w\tau} \bar{w}_k \bar{\tau}_c \\ &\quad + \sum_{c=1}^C \lambda_c^{pL^*} \bar{p}_c L^* + \sum_{c=1}^C \lambda_c^{\tau L^*} \bar{\tau}_c L^* + \sum_{k=1}^{K-1} \lambda_k^{wL^*} \bar{w}_k L^*, \end{aligned} \tag{3}$$

where  $\bar{\Pi} = \frac{\Pi}{w_K}$ ,  $\bar{p}_c = \frac{p_c}{w_K}$ ,  $\bar{w}_k = \frac{w_k}{w_K}$ ,  $\bar{\tau}_c = \frac{\tau_c}{w_K}$  indicate, respectively, normalized profit, normalized output price and normalized subsidy rate, and input price  $w_K$  is chosen as numeraire.

Differentiating the profit in (3) with respect to output prices  $\bar{p}_c$  yields the output level of crop  $c$  (Hotelling Lemma):

$$\begin{aligned} y_c = \frac{\partial \bar{\Pi}}{\partial \bar{p}_c} &= \alpha_c + \sum_{c'=1}^C \alpha_{cc'} \bar{p}_{c'} + \sum_{k=1}^{K-1} \delta_{ck}^{pw} \bar{w}_k + \sum_{c'=1}^C \delta_{cc'}^{p\tau} \bar{\tau}_{c'} \\ &\quad + \lambda_c^{pL^*} L^*, \forall c = 1, \dots, C. \end{aligned} \tag{4}$$

Differentiating the profit in (3) with respect to input prices  $\bar{w}_k$  yields the variable input demand equation (Hotelling Lemma):

$$\begin{aligned} -x_k = \frac{\partial \bar{\Pi}}{\partial \bar{w}_k} &= \beta_k + \sum_{k'=1}^{K-1} \beta_{kk'} \bar{w}_{k'} + \sum_{c=1}^C \delta_{ck}^{pw} \bar{p}_c + \sum_{c=1}^C \delta_{cc'}^{w\tau} \bar{\tau}_{c'} \\ &\quad + \lambda_k^{wL^*} L^*, \forall k = 1, \dots, K - 1. \end{aligned} \tag{5}$$

Differentiating profit in (3) with respect to subsidy rates  $\bar{\tau}_c$  yields the land allocation equation:

$$(1 - \phi) * l_c = \frac{\partial \Pi}{\partial \bar{\tau}_c} = \gamma_c + \sum_{c'=1}^C \gamma_{cc'} \bar{p}_{c'} + \sum_{k=1}^{K-1} \delta_{ck}^{w\tau} \bar{w}_k + \sum_{c'=1}^C \delta_{cc'}^{p\tau} \bar{\tau}_c + \lambda_c^{\tau L^*} L^*, \forall c = 1, \dots, C. \tag{6}$$

Profit function properties imply that the profit function is (i) non-decreasing in output prices  $p$ , non-increasing in input prices  $w$ , (ii) homogeneous of degree 1 in prices  $(p, w)$ , (iii) convex in prices  $(p, w)$ , and (iv) continuous in prices  $(p, w)$ . These properties imply some conditions to impose on the parameters. With the normalized form of the profit, the condition of linear homogeneity is automatically satisfied. The convexity conditions imply that the Hessian matrix is symmetrical and positively semi-definite. Imposing convexity restrictions is equivalent to imposing positive semi-definiteness on the matrix of parameters.<sup>6</sup>

Another regularity condition is the land adding-up condition  $\sum_{c=1}^C l_c = L^*$ , which implies the following conditions on the parameters:

$$\sum_{c=1}^C \gamma_{cc'} = \sum_{c=1}^C \delta_{ck}^{w\tau} = \sum_{c=1}^C \delta_{cc'}^{p\tau} = \sum_{c=1}^C \gamma_c = 0; \forall k, \forall c, \tag{7}$$

$$\sum_{c=1}^C \lambda_c^{\tau L^*} = 1. \tag{8}$$

The model to be estimated consists of the system of Eqs. (3, 4, 5, 6) after imposing convexity restrictions and land adding-up conditions 7–9. We let  $s_j(\theta)$  denote the  $j$ -th structural equation in the system, depending on exogenous covariates and a vector of parameters  $\theta$ , where the number of equations depends on the number of inputs, outputs and land use. In order to obtain precise results for policy analysis, we propose in this paper the estimation of this system of equations while explicitly dealing with corner solutions and the panel structure of our data. Contrary to other applications of land-use models in agriculture [3], the system of output, input and land equations depends only on observed prices and subsidies, and on total crop land. Other applications of land-use models consider input and output equations as an explicit function of land percentages for crops. We do not follow this approach here, as the demand for crop land is part of our structural system, which contains only exogenous covariates (more precisely, from the farmer’s point of view, assuming total crop land is fixed).

Imposing regularity conditions, as discussed above, is not the only issue when estimating a system of equations derived from profit maximization with farm-level data. Another concern is the existence of corner solutions, that is, zero land area and output level for some crops, and possibly zero expenditure on some inputs. As farmers rarely consider the same cropping system for every agricultural season because of agronomic and pest-management considerations, all possible crops are not planted every year by a particular farmer, implying that land and production variables are equal to 0 in this case. It is beyond the scope of this paper to provide a structural representation of cropping rotations through a dynamic model, see, e.g., [5, 48, 49]. Nevertheless, we provide below an original solution to this problem by considering a multivariate selection problem and relaxing some assumptions underlying the usual multivariate Tobit model.

### 3.2 Estimation Issues

#### 3.2.1 Corner Solutions

Estimating models with multivariate selection often implies a trade-off between computer-intensive numerical procedures and strong distributional assumptions to achieve parameter consistency and efficiency. For example, [2] discuss solutions based upon the [34] approach, which requires normality and homoscedasticity of structural error terms.

Our strategy is based on a multivariate version of the Tobit model for censored equations and, contrary to [5], we consider that prices and subsidies jointly determine the probability of a crop and its associated output level and land use (as would be the case in the original Tobit model). A drawback with the procedure described in [5] is that their estimator is consistent only if selection equations are uncorrelated, conditional on a the set of covariates. Although this condition can be tested in practice, this may limit the scope of the method. Furthermore, the correlation pattern between structural and selection equation is also restricted to a linear form, which may depend, however, on the period, as in [50]. In order to allow for correlation between all structural equations in the case of a multivariate Tobit model, one possibility is to consider the QML approach, based on pairwise joint probabilities of equations.<sup>7</sup> This is a simple alternative to multiple integration, which consists of exploring potential correlation between all pairs of structural equations.

Let  $s_{ij}(\theta)$  denote observation  $i$  of the  $j$ -th structural equation, evaluated at parameter vector  $\theta$ . The residual of equation  $j$  is denoted  $h_{ij}(\theta) = Z_{ij} - s_{ij}(\theta)$ , where  $Z_{ij}$  is the dependent variable in equation  $j$ .

<sup>6</sup> Imposing positive semi-definiteness on a matrix  $B$  is equivalent to writing  $B = AA'$  where  $A$  is a lower triangular matrix of the same dimension as  $B$ .

<sup>7</sup> A parametric version of this estimator has been proposed by [2] and applied to land-use decisions by [3].

The bivariate likelihood for observation  $i$  is given by

$$\mathcal{L}_i = \prod_{j=2}^{N-1} \prod_{k=1}^{j-1} \mathcal{L}_{ijk}, \quad (9)$$

where  $N$  is the total number of observations, and

$$\begin{aligned} \mathcal{L}_{ijk} &= \{F_2(h_{ij}, h_{ik})\}^{1(Z_{ij}=0, Z_{ik}=0)} \\ &\times \{f_2(h_{ij}, h_{ik})\}^{1(Z_{ij}>0, Z_{ik}>0)} \\ &\times \{F(h_{ik}|h_{ij}) \times f(h_{ij})\}^{1(Z_{ij}>0, Z_{ik}=0)} \\ &\times \{F(h_{ij}|h_{ik}) \times f(h_{ik})\}^{1(Z_{ij}=0, Z_{ik}>0)}. \end{aligned} \quad (10)$$

Denote  $h_{ij}^*(\theta) = [Z_{ij} - s_{ij}(\theta)]/\sigma_j$  the standardized residual of equation  $j$ , where  $Z_{ij}$  is the dependent variable and  $\sigma_j$  the standard deviation of the residual in equation  $j$ .  $F$  is the cumulative distribution function and  $f$  is the density function of  $h_{ij}^*(\theta)$ .

Under the normality assumption, the individual contribution to the likelihood becomes

$$\begin{aligned} \mathcal{L}_{ijk} &= \left\{ \Psi(h_{ij}^*, h_{ik}^*, \rho_{jk}) \right\}^{1(Z_{ij}=0, Z_{ik}=0)} \\ &\times \left\{ \psi(h_{ij}^*, h_{ik}^*, \rho_{jk}) / (\sigma_j \sigma_k) \right\}^{1(Z_{ij}>0, Z_{ik}>0)} \\ &\times \left\{ \frac{\phi(h_{ij}^*)}{\sigma_j} \Phi \left[ \frac{h_{ik}^* - \rho_{jk} h_{ij}^*}{\sqrt{1 - \rho_{jk}^2}} \right] \right\}^{1(Z_{ij}>0, Z_{ik}=0)} \\ &\times \left\{ \frac{\phi(h_{ik}^*)}{\sigma_k} \Phi \left[ \frac{h_{ij}^* - \rho_{jk} h_{ik}^*}{\sqrt{1 - \rho_{jk}^2}} \right] \right\}^{1(Z_{ij}=0, Z_{ik}>0)}, \end{aligned} \quad (11)$$

where  $\Psi(\cdot, \cdot, \cdot)$  and  $\phi(\cdot, \cdot, \cdot)$  are the bivariate cumulative density and density functions, respectively;  $\phi(\cdot)$  and  $\Phi(\cdot)$ , respectively, denote the univariate density and cumulative density functions of the standard Normal distribution. Maximizing (10) with respect to  $\theta$  and the variance–covariance matrix yields consistent quasi-maximum likelihood (QML) estimates. As discussed in [4] and in [51], the QML estimator is consistent and asymptotically normal if and conditional mean and variance structure of the model are correctly specified, and if the log-likelihood function used for QML belongs to the generalized linear exponential family.<sup>8</sup>

### 3.2.2 The Semi-nonparametric Estimator

As discussed above, pairwise computation of distribution functions is much more straightforward than full-information

estimators such as maximum likelihood based on all (potentially censored) structural equations. Although QML estimation as described above is a convenient alternative to maximum likelihood under multivariate censoring, it has several drawbacks. First, efficiency of QML estimates may depend on the criterion used in practice (in the family of linear exponential distributions, see [4]) in the maximization of the log-likelihood. Second, as a corollary to the conditions above, [38] and [2] acknowledge that the QML estimator above is not robust to deviations from homoscedasticity assumptions. To deal with the first issue, an interesting alternative is a flexible estimation method, such as a semiparametric alternative to the parametric QML.<sup>9</sup> Semiparametric QML estimation of univariate Tobit-censored models was proposed as early as [52], and their asymptotic properties are discussed in, e.g., [6]. In semiparametric estimation, normal density and cumulative density functions are replaced in [2]’s QML objective by nonparametric approximations. The distribution of equation residuals is then left unrestricted and structural parameters  $\theta$  can be estimated jointly with nonparametric estimation of the univariate and bivariate distributions.

The semiparametric QML may also provide a solution to the second issue, because variance–covariance parameters can be considered nuisance parameters and be replaced by data-driven bandwidth parameters in, e.g., kernel approximation of distribution functions. By doing so, semiparametric estimation may significantly reduce the number of parameters to estimate, but it may be computer-intensive and have a lower convergence rate. Another alternative, which we consider in this paper, is to consider semi-nonparametric estimation instead.<sup>10</sup> The approach behind our semi-nonparametric model with selection is to replace unknown joint distributions by a series approximation (see [53]), while providing a solution to the fact that such approximation may cause the number of terms to grow with the sample size [54].

This solution uses *sieve estimation*, i.e., replacing the infinite-dimensional space associated with the nonparametric functions by a flexible parametric one. Conditions for consistency and asymptotic normality of sieve ML estimation have been provided by [55] and [1].

Inspecting the form of the QML estimator (11), we see that four functions need to be estimated: a) the bivariate cumulative density function,  $F_2(h_{ij}, h_{ik})$ ; b) the bivariate density

<sup>8</sup> Or only to the linear exponential family if only the conditional mean is specified, see [4].

<sup>9</sup> A more flexible estimator is in general less efficient if the restrictions implied by the less flexible estimator are valid, such as, e.g., normality or homoscedasticity.

<sup>10</sup> Following [1], a model is called parametric if its parameters are in finite-dimensional spaces, nonparametric if its parameters are in infinite-dimensional spaces, semiparametric if its parameters of interest are in finite-dimensional spaces and its nuisance parameters are in infinite-dimensional spaces, and semi-nonparametric if it contains both finite-dimensional and infinite-dimensional unknown parameters of interest.



function,  $f_2(h_{ij}, h_{ik})$ ; c) the univariate conditional cumulative density function,  $F(h_{ij}|h_{ik})$ ; d) the univariate density function,  $f(h_{ik})$ . The bivariate density function can be computed from the derivative of an approximation to the bivariate cumulative distribution function, obtained from a tensor-product spline, such tensor-product spline approximations being easy to compute from a B-spline representation [55].

An interesting aspect of B-splines is that they can be constructed under some shape and smoothness restrictions that are easy to impose, as well as their derivatives of any order, which provide us with a natural procedure for estimating probability and cumulative density functions:

- Compute the empirical cdfs in the univariate and bivariate cases, for each selection regime;
- Approximate the univariate and bivariate cdfs for every selection regime by B-splines, imposing smoothness and non-decreasing approximations over the [0, 1] interval;
- Compute the derivative of the spline approximation to obtain the pdf;
- Combine joint and marginal distributions to obtain the conditional cdfs for each selection regime.

An important aspect of our semi-nonparametric procedure is that variance and covariance parameters are not estimated jointly with structural parameters, but are replaced by data-driven dispersion measures. The latter are equivalent to bandwidth parameters in nonparametric estimation and correspond to normalization parameters in sieve (B splines) estimation. Hence, they are considered nuisance parameters and homoscedasticity or other restrictions on variances and covariances are completely avoided, as opposed to the parametric QML estimation.

### 3.2.3 Panel Data

Equations (4), (5) and (6) in Sect. 3 have an implicit static form that can accommodate panel data, by simply adding a time index,  $t$ , to dependent and explanatory variables. This also applies to residuals from structural equations that become  $s_{ijt}(\theta)$  for farmer  $i$ , equation  $j$  and time period  $t$ .

To consider QML estimation, whether parametric or semi-nonparametric, an important condition is that sample observations are independently and identically distributed (i.i.d.). With panel data however, this is obviously not the case because unobserved individual effects are likely to be present. Model estimation procedures discussed above have therefore to be adapted to the case of panel data. Consider the total number of observations  $N = \sum_{i=1}^n T_i$ , where  $n$  is the number of individuals (cross-sectional units), and  $T_i$  denotes observations for cross-sectional unit  $i$ . In other words, unbalanced panels can be accommodated for, when  $T_i$  is different across cross sections.

To control for unobserved individual heterogeneity, possibly correlated with explanatory variables, we may consider a fixed-effect approach to the production model. However, as the above model is nonlinear, within-type estimators would not be consistent with a fixed number of time periods. We choose to control for such unobserved heterogeneity by implementing the Mundlak method (see [50]). Assume

$$h_{ijt}(\theta) = Z_{ijt} - s_{ijt}(\theta) - \eta_{ij}, \quad i = 1, 2, \dots, N,$$

where  $\eta_{ij}$  is the individual effect in equation  $j$ , possibly correlated with explanatory variables in  $s_{ijt}(\theta)$ . Consider first the balanced panel data case, with  $T_i = T, \forall i$ . We further assume

$$\eta_{ij} = X_{ij1}\gamma_{j1} + \dots + X_{ijT}\gamma_{jT} + v_{ij},$$

where  $X_{ijt}$  is the  $K$ -vector of explanatory variables in structural equation  $s_{ijt}(\theta)$ , such that  $E(v_{ij}|X_{ijt}) = 0$ . Under this conditional moment condition,  $\eta_{ij}$  can be replaced in the structural equation by its projection onto explanatory variables. The Mundlak approach corresponds to the special case of [56] with  $\gamma_{j1} = \dots = \gamma_{jT} = \bar{\gamma}_j/T, \forall j$  and is preferred in practice if the number of explanatory variables is large, because it inflates the number of parameters by  $K \times T$ . Such conditioning is typically designed for discrete-choice models (Probit, Logit) and is naturally adapted to the Tobit framework. Because unobserved heterogeneity in structural and selection equations is likely to be correlated with explanatory variables, we do not consider random-effects estimation but only fixed-effects estimation, to control such possible source of bias in parameter estimates.

The Mundlak approach for fixed effects is easily adapted to the unbalanced panel-data case, with

$$\eta_{ij} = \left( \frac{1}{T_i} \sum_{t=1}^{T_i} X_{ijt} \right) \gamma_j + v_{ij},$$

where  $T_i$  is the individual-specific number of periods. With large  $N$ , we can consistently estimate parameters  $\gamma_j$  even when the panel data set is unbalanced.

When proceeding with estimation, we make the important assumption that, conditional on fixed effects  $\eta_{ij}$ , the error terms are i.i.d., so that the QML framework for the parametric and the semi-nonparametric estimators applies.

## 4 Data and Econometric Estimation Results

### 4.1 Data

The empirical application is conducted on a sample of French farmers from the *département* of Meuse. The data were provided by the *Centre d'Economie Rurale de La Meuse*, an agricultural extension service that provides farmers with assistance

in bookkeeping and auditing. The original sample consists of 2,356 farm-year observations on 638 farmers observed between 2006 and 2010. This sample represents about 21 percent of the total number of farms in this *département* (2,975 in 2010, of which about 2,100 are classified as medium- to large-sized farms, see [57]). We remove yearly observations corresponding to no-crop production (animal production only, or no agricultural production for a particular year) or no land use for any of the following: wheat, barley, rapeseed, diester and non-compulsory set-aside. The latter selection restricts the sample to consider only major groups of crops cultivated in the *Meuse* département, discarding specialty and some industrial crops that are much less represented in the region, and which are not always associated with CAP area-based payments. The final sample consists of 2,014 farm-year observations on 524 farmers over five years. The final dataset represents about 56 percent of specialized arable farms in the *Meuse* département in 2010 (927 farms within this category, according to [57]). An interesting feature of our data and the period considered is that, during that period, France operated a decoupling scheme with a hybrid status. Following the European CAP reform in 2003 (Luxembourg Agreement, see [58]) decoupling was introduced with environmental cross-compliance among other measures, aimed at making farmers' production decisions more market-oriented [59]. Member States retained some flexibility in the choice between partial and full decoupling. For example, Spain, France and Portugal opted to maintain the maximum permitted amount of coupled payments (25%), in both the livestock and arable crop sectors. In the French case, the 2008 CAP health check provided for an "à la carte" selection of the tools, allowing voluntary implementation to start by 2010 and end by 2012 at the latest [60]. In terms of the model notation in subsection 3.1, the proportion of area-based subsidies that is decoupled from production,  $\phi$ , is equal to 0.75. Another interesting feature of our sample is the higher variability in area-based unit payment rates over the 2006–2010 period, compared with the subsequent years. Such variability is necessary to identify subsidy effects alongside price effects in the structural equations, as partial coupling was also necessary to identify crop-specific output price effects.

#### 4.1.1 Sample Description

The final dataset is unbalanced in the following way: farmers present from 2006 to 2010: 49.8 percent; farmers present four years out of five: 15.45 percent; three years out of five: 10.69 percent; two years out of five: 7.82 percent; one year only: 7.63 percent. In terms of yearly observations, 417 farmers were

observed during year 2006, 420 in year 2007, 422 in year 2008, 386 in 2009 and 369 in year 2010.<sup>11</sup> The average total farm area is 197.9 ha, of which 121.41 ha of arable land (standard deviation of 75.65 ha) and 74.9 for permanent grazing and temporary land for pasture. Average production cost is 263,777.82 Euros / year and total profit per farm is 56,783.74 Euros per year, about 285.18 Euros per ha (standard deviation of 598.82). These statistics of dispersion indicate that farm diversity is limited, as far as size and economic performance are concerned. In terms of spatial location, farms are widespread over the whole *Meuse* département, as can be seen from Fig. 1, which represents the spatial distribution of sample farms in the *cantons* (an administrative unit, between the *commune* and the *département*). The total area covered by the sample is relatively limited (about 6,200 square kilometers, about 2,400 sq. miles), so that differences in climate and soil characteristics are fairly limited as well. Moreover, as fixed effect procedures are employed in the estimation, farm-specific or site-specific non-time-varying characteristics will be filtered out from the model.

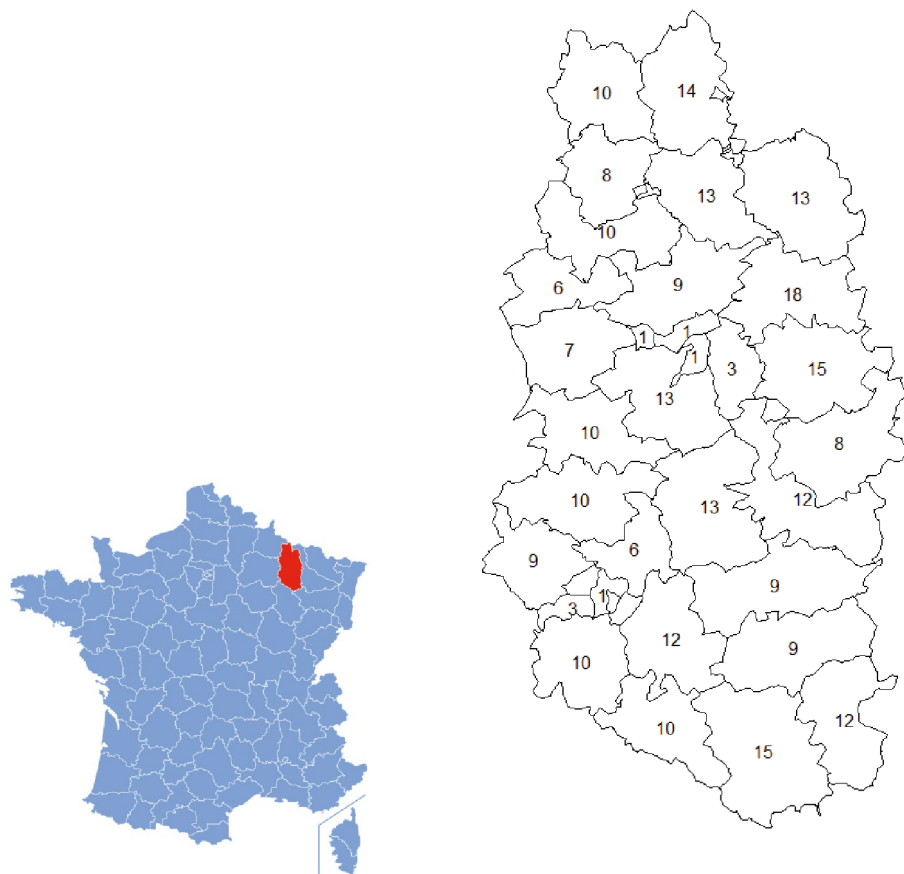
Concerning crops and inputs, we use the major cropping systems in the *Meuse* département and select wheat, barley, rapeseed and diester, the last of which is used for biofuel production. Because wheat is, in the vast majority of cases, associated with barley, and farm-gate prices of both crops move in parallel, we combine them to form a composite cereal output. Land decisions are associated with these crops (cereals, rapeseed and diester) as well as voluntary set-aside, as discussed above (it should be recalled that set-aside obligations under the CAP are already accounted for in the model, as total land  $L^*$  considered in the estimation is net of the compulsory set-aside area). We consider only three inputs: seed, fertilizer and pesticide, which are considered the most crop-specific and therefore whose demand is more likely to be influenced by a change in cropping pattern (for our set of arable crops, as opposed to, e.g., labor and fuel).<sup>12</sup> To obtain farm-specific input prices that vary over time, we proceed as follows. As database records contain only input expenditures and not input physical quantities nor unit prices, we use official yearly statistics on agricultural input price indexes at the département level. We then convert to farm-specific price indexes using the Tornqvist formula, with 2004 as the baseline year.<sup>13</sup> We then check for possible multicollinearity in input and output prices, with

<sup>11</sup> The imbalanced pattern is somewhat different to that of the original sample, where 39.50 percent of farmers were observed over the entire period, 21.31 percent over four years out of five, 9.56 percent over three years, 4.39 percent over two years and 7.99 percent over one year.

<sup>12</sup> In our sample, fertilizer and pesticide account on average for, respectively, 9.49 and 8.25 percent of operating costs, compared with 9.06 percent for wage labor, 4.18 percent for fuel and 3.73 percent for seed.

<sup>13</sup> The Tornqvist price index of input  $j$  for farm  $i$  and year  $t$  is computed as follows:  $\log p_{ijt} = (w_{ijt} + w_{j0}) \log(p_{jt}/p_{j0})$ , where  $w_{ijt}$  and  $w_{j0}$ , respectively, denote the percentage cost of inputs  $j$  for farm  $i$  at year  $t$  and for baseline year 0;  $p_{jt}$  and  $p_{j0}$  denote the unit input price of input  $j$  for year  $t$  and for baseline period 0, respectively.

**Fig. 1** Location of sample farms in *cantons* of the Meuse *département*. Note. Numbers for each *canton* represent the total number of farms in the sample



Note. Numbers for each *canton* represent the total number of farms in the sample.

the condition number, computed every year and over the whole sample. The condition number is always less than 15, confirming that multicollinearity in input prices is limited.<sup>14</sup> Crop outputs are in tons, fertilizer input is in kg, and pesticide input is computed from pesticide expenditure divided by its price index.

All prices and unit subsidies are normalized to the unit cost of seed so that in the estimation we consider only the other two inputs.<sup>15</sup>

<sup>14</sup> Note that the geographical area of our study is relatively limited, so that variability of input prices may seem to be artificially augmented by the use of the Tornqvist price index. There is, however, an important aspect to consider, even with such limited geographical area: most farm production inputs have some degree of heterogeneity in their characteristics (quality, in a broad sense). Even chemical (mineral) fertilizer has a different combination of nitrogen, phosphorous, potash, urea, etc., and pesticide input is even more heterogeneous in its composition. This means that variability in input price reflects such differences in “quality” at the farm level. Moreover, there may be different unit prices paid by farmers, depending on quantity purchased from retailers, which are transmitted to percentages of farm-level costs, leading to different price index levels even for similar input quantities.

Our sample covers the pre- and the post-2008 period, where compulsory set-aside was abandoned and industrial crops such as biofuels were allowed on set-aside land with the same subsidy rate as other arable crops (different from the agronomic set-aside subsidy rate). This concerns, in our case, rapeseed that can be produced for energy use or for the agrofood industry, with the same area-based subsidy rate after 2008. Our data at the farm and crop level allow us to account for the different cases, i.e., area under rapeseed for industrial use has a different area-based unit payment before and after 2008, while it remains at the same rate for rapeseed sold to the food industry. Because changes in use-specific area-based payments over the period are fully accounted for at the farm level when constructing our payment  $\tau_c$  and area  $l_c$  variables, we assume that the model parameters will not depend on such policy changes. Therefore, we do not include

<sup>15</sup> In order to check how our results are sensitive to their choice of numeraire, we have estimated the model with alternative options for the numeraire, including, e.g., cereal output price (weighted average of wheat and barley) or pesticide price. Estimation results were not different in terms of parameter signs for price and subsidy effects, and only slight differences were found in the magnitude of elasticities.

**Table 2** Descriptive statistics

Variable	Mean	Std. dev.	Proportion > 0
Cereal output (tons)	532.3851	342.0180	0.9985
Rapeseed output (tons)	88.4597	78.2572	0.8332
Diester output (tons)	11.0806	20.9001	0.3133
Fertilizer input (kg)	590.4984	200.7343	0.9911
Pesticide input	163.5981	118.3055	0.9916
Cereal area (ha)	84.5566	53.7241	0.9995
Rapeseed area (ha)	26.2255	5.4881	22.90
Diester area (ha)	3.4394	6.3397	0.3153
Land set-aside (ha)	4.1273	6.5696	0.7066
Cereal price	1.0207	0.3177	
Rapeseed price	2.2857	0.5206	
Diester price	1.2061	0.9199	
Fertilizer price	0.2756	0.1274	
Pesticide price	0.9552	0.0306	
Cereal subsidy	5.4009	2.7268	
Rapeseed subsidy	5.9490	1.3463	
Diester subsidy	5.1049	2.7187	
Land set-aside subsidy	8.6846	2.8985	

2014 observations. Price and subsidy variables (in Euros) are normalized by the unit price of seed. Crop outputs are in tons, fertilizer input is in kg, and pesticide input is computed from pesticide expenditure divided by its price index. Cereal is the combination of wheat and barley crops. Land set-aside corresponds to the area above the non-compulsory set-aside requirement (see text)

in the model a dummy variable for the pre- and post-2008 period, as we assume parameters are neither period- nor policy-dependent. Moreover, as mentioned in the data subsection above, the eligibility condition for payments under CAP was a minimum arable area of 0.3 ha. No farmers in our sample failed to meet this criterion, so that all were eligible and actually received some form of area-based CAP payments every year from 2006 to 2010.

Table 2 presents descriptive statistics for the sample. Almost all farmers grow cereals (wheat and barley) while the respective proportions of positive land percentages for rapeseed, diester and land set-aside are 83%, 31% and 70% of the full sample. These descriptive statistics need to be interpreted with caution since some farmers may not grow a particular crop over the whole period.

#### 4.1.2 Intensity of Input Indicators

As discussed above, one reason for using detailed farm-level and crop-level production data is that input use (and therefore, environmental indicators) may differ across cropping systems, implying that farmers' decisions including corner solutions should be modeled explicitly. Using the sample of farmers, we compute average environmental indicators for various cropping systems: cereals only, cereals-rapeseed,

**Table 3** Environmental indicators, by crop combination

Cropping system	<i>N</i>	Indicator			
		$x_F$	$x_P$	$II_F$	$II_P$
<i>C</i> only	312	270.77 (118.13)	34.52 (35.40)	17.48 (7.04)	1.82 (1.06)
<i>C + R</i> only	1067	585.08 (111.64)	188.13 (114.21)	17.01 (6.43)	4.52 (1.27)
<i>C + D</i> only	21	516.06 (122.38)	100.86 (74.33)	24.61 (7.05)	4.13 (1.64)
<i>C + R + D</i>	610	767.75 (136.25)	189.44 (110.05)	22.79 (8.41)	4.72 (1.37)

*N* denotes the number of observations.  $x_F$ ,  $x_P$ ,  $II_F$  and  $II_P$ , respectively, denote input demand for fertilizer and pesticide, and intensity of input use for fertilizer and pesticide. Cropping systems are defined as combinations of cereals (*C*), rapeseed (*R*) and diester (*D*). Standard deviations are in parentheses

etc.  $II_k$  denotes the intensity of input *k* (*k* = fertilizer, pesticide) demand on cultivated land, defined as

$$II_k = \frac{\text{input } k \text{ demand}}{\text{cropland}} = \frac{x_k}{(l_{\text{cereal}} + l_{\text{rapeseed}} + l_{\text{diester}})} = \frac{x_k}{L - l_{\text{setaside}}} \quad (12)$$

Descriptive statistics for our environmental indicators are presented in Table 3. Means and standard deviations are computed for a given combination of crops (e.g., cereals-diester) over all corresponding observations across farmers and years. Crop combinations with fewer than 10 observations are discarded. Table 3 confirms that input use is strongly heterogeneous across cropping systems. This results in very different average environmental indicators and implies that accounting for farmers' decisions over a whole cropping system (i.e., including decisions leading to corner solutions for some crops) is preferable in terms of policy impact evaluation. Cereal-only cropping systems (*C*) have a lower fertilizer and pesticide input use ( $x_F$  and  $x_P$ ), and a lower input intensity indicator, except for fertilizer where  $II_F$  is slightly higher for cereals only than for cereal and rapeseed (*C + R* only). The cereal-rapeseed (*C + R*) crop combination is associated with a fertilizer input intensity very close to cereal alone (*C*), while pesticide input intensity is much higher in the (*C + R*) system than in *C*. This is presumably due to the fact that the total area of the cereals cropping system is smaller than the area of the cereals plus rapeseed cropping system, while intensities of input use for fertilizer are comparable. The difference in input intensity is more pronounced, especially for fertilizer, when diester is included in the crop combination (*C + D* or *C + R + D*). Interestingly, cropping systems that involve diester (*C + R + D*) have a lower fertilizer use intensity than *C + D* only, but a higher intensity of pesticide use (4.72 compared with 1.82, 4.52 and 4.13).

**Table 4** Goodness-of-fit measures for parametric and semi-nonparametric estimators

Structural Equation	Estimation Method	R (sample) <sup>2</sup>	R <sup>2</sup> (obs.> 0)	HR Test Statistic	Correct Predictions (percent)
<i>q</i> cereal	PQML	0.9157	0.9192	-0.90 (0.36)	99.85
	SNPQML	0.9102	0.9144	0.92 (0.35)	99.85
<i>q</i> rapeseed	PQML	0.7661	0.7369	0.46 (0.65)	83.32
	SNPQML	0.7544	0.7346	0.30 (0.76)	83.12
<i>q</i> diester	PQML	0.5093	0.3193	0.78 (0.43)	79.69
	SNPQML	0.4101	0.4243	1.61 (0.10)	82.22
<i>x</i> ferti.	PQML	0.5225	0.5445	0.61 (0.54)	99.26
	SNPQML	0.4147	0.4331	0.71 (0.47)	99.60
<i>x</i> pest.	PQML	0.8304	0.8376	2.47 (0.01)	99.16
	SNPQML	0.8309	0.8383	2.72 (0.00)	99.16
<i>l</i> cereal	PQML	0.9617	0.9607	0.25 (0.80)	99.95
	SNPQML	0.9617	0.9616	1.31 (0.19)	99.95
<i>l</i> rapeseed	PQML	0.8195	0.8001	1.00 (0.31)	83.42
	SNPQML	0.8192	0.8075	0.83 (0.40)	83.42
<i>l</i> diester	PQML	0.5288	0.3445	0.34 (0.73)	79.15
	SNPQML	0.5123	0.4332	1.98 (0.05)	80.73
<i>l</i> set-aside	PQML	0.1232	0.1337	-2.13 (0.03)	72.54
	SNPQML	0.1248	0.1066	-1.75 (0.08)	77.26

2014 observations. PQML and SNPQML, respectively, denote parametric and semi-nonparametric quasi-maximum likelihood estimation. HR test statistic (*p*-value in parentheses) is Born and Breitung (2016)'s Heteroskedasticity-Robust test for serial correlation in panel data. The proportion of correct predictions is the sum of sensitivity (correctly predicted positive outcomes) and specificity (correctly predicted negative outcomes) measures. The proportion of positive outcomes (resp., negative) wrongly predicted as negative (resp., positive) outcomes is computed with respect to the number of positive (resp., negative) outcomes. The last column gives goodness-of-fit measurements designed for discrete outcomes computed from estimated probabilities

## 4.2 Empirical Results

### 4.2.1 Relative Performance of Parametric and Semi-nonparametric Methods

To evaluate the relative performance of the parametric and semi-nonparametric QML estimators with respect to our panel data sample, we compute goodness-of-fit measures for the continuous and discrete parts of the model. We first compute R<sup>2</sup>s on the full sample, for each structural equation, without accounting for the contribution of estimated fixed effects. Because such goodness-of-fit measures are difficult to interpret when the proportion of zero observations is large, we also produce R<sup>2</sup>s for each equation on the subset of positive observations only. Columns 2 and 3 of Table 4 present computed R<sup>2</sup>s for the parametric and semi-nonparametric QML estimators. We note that the difference between R<sup>2</sup>s on the full sample or the subsample of positive observations is noticeable mostly when the proportion of censored observations is higher (case of *q* rapeseed, *q* diester, *l* rapeseed, *l* diester and *l* set-aside), with the R<sup>2</sup>s on the full sample being lower than

the restricted version in 12 cases out of 27. Interestingly, the model fits data better with parametric PQML than with semi-nonparametric SNPQML on the full sample (6 cases out of 9), but this is the opposite on the restricted samples. The semi-nonparametric estimator performs better than PQML on subsamples with positive observations only.

To check for serial correlation in residuals estimated from PQML and SNPQML, we compute the heteroskedasticity-robust test statistic (HR) proposed by [61] in the context of a panel data model with fixed effects. Test results reported in Table 4 indicate that serial correlation is not present in 15 cases out of 18, with a 5-percent significance level, so that the static specification is valid in a majority of structural equations.

Turning now to the goodness-of-fit measures for discrete outcomes (last column of Table 4), it is more interesting to focus on the equations with a significant proposition of zero observations (*q* rapeseed, *q* diester, *l* rapeseed, *l* diester and *l* set-aside). For these five equations, the proportion of correct predictions (positive and negative outcomes) is fairly similar for PQML and SNPQML, so that the gain associated with the latter is only minor.

**Table 5** Parametric QML elasticity estimates with fixed effects

	$p$ cereal	$p$ rapeseed	$p$ diester	$w$ fert.	$w$ pest.	$\tau$ cereal	$\tau$ rapeseed	$\tau$ diester	$\tau$ set-aside
$p$ cereal	0.0096*** (0.0009)	0.0353*** (0.0015)	-0.0707*** (0.0035)	-0.0117*** (0.0008)	0.0120*** (0.0006)	0.0116*** (0.0012)	0.0239*** (0.0011)	-0.0790*** (0.0037)	0.0259*** (0.0013)
$p$ rapeseed		0.1418*** (0.0042)	-0.3020*** (0.0048)	-0.0555*** (0.0010)	0.0548*** (0.0014)	0.0280*** (0.0034)	0.0883*** (0.0009)	-0.3356*** (0.0059)	0.1249*** (0.0020)
$p$ diester			0.6714*** (0.0023)	0.1261*** (0.0028)	-0.0958*** (0.0034)	-0.0293*** (0.0058)	-0.1760*** (0.0006)	0.7460*** (0.0014)	-0.2987*** (0.0046)
$w$ fert.				0.0507*** (0.0013)	-0.0394*** (0.0026)	0.0010 (0.0014)	-0.0350*** (0.0012)	0.1200*** (0.0030)	-0.0589*** (0.0018)
$w$ pest.					0.3700*** (0.0032)	0.0498*** (0.0037)	0.0316*** (0.0037)	-0.1071*** (0.0020)	-0.0290*** (0.0026)
$\tau$ cereal						0.0537*** (0.0017)	0.0022 (0.0027)	-0.0448*** (0.0044)	-0.0104*** (0.0015)
$\tau$ rapeseed							0.1475*** (0.0011)	-0.2180*** (0.0014)	0.0638*** (0.0011)
$\tau$ diester								0.9046*** (0.0018)	-0.3314*** (0.0054)
$\tau$ set-aside									0.1479*** (0.0044)

2014 observations.  $p$ ,  $w$  and  $\tau$ , respectively, denote output prices, input prices and area-based subsidy rates, and fert. and pest. denote fertilizer and pesticide inputs. Robust standard errors are in parentheses. Dependent variables are in rows; price/subsidy in columns. Elasticities of inputs are computed with the minus sign

\*, \*\* and \*\*\*, respectively, denote parameter significance at the 10, 5 and 1 percent level

#### 4.2.2 Estimation Results

Results of the estimated elasticities for the parametric and semi-nonparametric estimator are reported in Tables 5 and 6, respectively. The elasticity magnitude using SNPQML is similar to the estimates obtained using the PQML model, except for the impact of the set-aside subsidy, which is less significant in the output and land equations in the SNPQML estimation. Moreover, the own-price elasticity of fertilizer demand is no longer significant. Recall that Table 5 presents minus the elasticities of fertilizer and pesticide demand with respect to set-aside subsidy; most of them are not significant, although their values are close to those estimated using the parametric model. The elasticities of fertilizer and pesticide demand with respect to set-aside subsidy are not significant, even though their values are very close to those estimated with the parametric model. To compare PQML and SNPQML model specifications, a feasible closeness specification test is proposed by [62], as it is based on likelihood or quasi-likelihood contributions that are directly available from our parametric and semi-nonparametric estimations. This test is useful to decide between two non-nested models (e.g., the same subsets of parameter sets but different functional forms, see [63]). The Vuong specification test statistic equals 1.9077 (p-value=0.0564), when testing the null that both specifications are equal (against

the alternative that the parametric model is preferable to the semi-nonparametric one). We therefore conclude that we do not reject the hypothesis that both model specifications are equivalent.

To summarize our results in terms of comparison of the PQML and SNPQML estimators, the semi-nonparametric QML is less efficient than the parametric version even if the parameter estimates are fairly close. However, the semi-nonparametric estimator performs at least as well in terms of predicting output, input and land-use decisions, as well as in predicting the probability of corner solutions (discrete outcomes).

In the next section, we will discuss both estimators for the elasticity estimation results in terms of environmental impacts in subsections 5.1 and 5.2, but we will consider only the most efficient parametric PQML estimation for the simulation exercise in subsection 5.3.

## 5 Environmental Impacts of Set-aside Subsidy

The actual efficiency of set-aside policies for conserving the environment and protecting farmland biodiversity is widely debated. Some studies show that set-aside land improves biodiversity [7], while other studies find that set-aside policies are inefficient for conservation purposes [8]. Our

**Table 6** Semi-nonparametric QML elasticity estimates with fixed effects

	$p$ cereal	$p$ rapeseed	$p$ diester	$w$ fert.	$w$ pest.	$\tau$ cereal	$\tau$ rapeseed	$\tau$ diester	$\tau$ set-aside
$p$ cereal	0.0100*** (0.0011)	0.0371*** (0.0046)	-0.0760*** (0.0118)	-0.0118 (0.0128)	0.0128 (0.0105)	0.0132 (0.0101)	0.0235 (0.0183)	-0.0773*** (0.0163)	0.0290** (0.0103)
$p$ rapeseed		0.1542*** (0.0402)	-0.3288*** (0.0633)	-0.0630 (0.0563)	0.0598 (0.0563)	0.0267 (0.0527)	0.0937 (0.0987)	-0.3406*** (0.0491)	0.1376*** (0.0320)
$p$ diester			0.7126*** (0.1891)	0.1403 (0.0908)	-0.1024 (0.0656)	-0.0341 (0.1093)	-0.1962 (0.1638)	0.7444*** (0.1957)	-0.3077** (0.1556)
$w$ fert.				0.0516 (0.0563)	-0.0355 (0.0604)	0.0085 (0.0193)	-0.0403 (0.0776)	0.1354 (0.1119)	-0.0655** (0.0292)
$w$ pest.					0.3746*** (0.1110)	0.0552 (0.0557)	0.0361 (0.0967)	-0.1090* (0.0574)	-0.0298 (0.0771)
$\tau$ cereal						0.0717 (0.0452)	-0.0075 (0.0223)	-0.0343 (0.1430)	-0.0059 (0.0369)
$\tau$ rapeseed							0.1404 (0.1057)	-0.2210 (0.1693)	0.0790* (0.0421)
$\tau$ diester								0.8452*** (0.1060)	-0.3266* (0.1367)
$\tau$ set-aside									0.1412** (0.0954)

2014 observations.  $p$ ,  $w$  and  $\tau$ , respectively, denote output prices, input prices and area-based subsidy rates, and fert. and pest. denote fertilizer and pesticide inputs. Robust standard errors are in parentheses. Dependent variables are in rows; price/subsidy in columns. Elasticities of inputs are computed with the minus sign

\*, \*\* and \*\*\*, respectively, denote parameter significance at the 10, 5 and 1 percent level

purpose here is not to provide an assessment of direct environmental consequences of set-aside policy implementation. Rather, we insist on the importance of an ex ante evaluation of land-use policies such as set-aside and greening of the CAP, in a context of increasing “social demand” for more sustainable agricultural practices.

In the broader context of the CAP reform and, more generally, of agricultural policies embedding environmental objectives, our simulation experiment aims to document the implementation issues for tax policies on nonpoint source emissions such as from pesticide and fertilizer inputs. In the context of regulating nonpoint source pollution from agricultural sources, indirect tax-subsidy policy is sometimes advocated [64], such as subsidizing less input-intensive crops or alternative land use (e.g., land set-aside). To address the possibility that subsidizing set-aside could worsen the environmental effects associated with chemical inputs, we propose two environmental indicators, which can be linked to a policy instrument such as a set-aside subsidy.<sup>16</sup> Note that, because we use only production data, these indicators will be “environmental effect proxies” and will tend to measure environmental “pressure” from production rather than an actual impact on the ecosystem. The

first indicator is the elasticity of input demand (fertilizer, pesticide) with respect to the set-aside subsidy rate, which measures the sensitivity of farm-level input demand to a change in the unit set-aside subsidy rate, all else remaining equal. Assuming total farm land is constant, this indicator is relevant at the farm level and depends indirectly on land set-aside and crop decisions. The drawback of using this indicator is that the intensive margin (i.e., increasing input intensity per unit of land) is relevant only if computed for the cropped area [65]. Therefore, we consider a second indicator based on intensity of input use per unit of cultivated land, which allows us to measure the intensification effect of changes in the set-aside subsidy<sup>17</sup>. Note that this environmental indicator, as a proxy based on input use data, cannot accommodate possibly non-linear effects of cropping practices and cropping systems on the environment, which would require more detailed data at the plot level. Moreover, depending on the local ecosystem features, spatial spillover effects are also likely to exist between

<sup>16</sup> We make the distinction between the “environmental indicators” introduced above, which are defined in levels and ratios, and the “sensitivity indicators” that are computed from them and are unit-free.

<sup>17</sup> We remind here that pesticides are very heterogeneous with regard to their toxicity and that as shown in [81], quantity-based pesticide indicators could fail to take into account important qualitative toxicity characteristics of pesticides. Other indicators that consider quantities of active ingredients sold, spraying frequency and index of load could be used instead.

land plots with intensification and another set of plots with land set-aside. Our analysis being at the farm level only, such environmental effects are not identified in our analysis. See, for example, [66] and [67].

We discuss below the results of the estimation of our first indicator, namely the chemical input demand elasticities, with respect to set-aside subsidy (subsection 5.1). Then, we present the measurement and the results of calculations of the second indicator namely the chemical input intensity elasticities with respect to the set-aside subsidy rate (subsection 5.2). Finally in subsection 5.3, we evaluate the environmental impacts of a set-aside policy, obtained as proxies from pesticide and fertilizer demand and input intensity elasticities with respect to land set-aside subsidy.

### 5.1 Chemical Input Demand Elasticities with Respect to Set-aside Subsidy

Elasticity estimates from parametric QML, presented in Table 5, show that the set-aside area is significantly sensitive to its unit set-aside subsidy (elasticity = 0.1479), an increase in the latter also implying an increase in output and planted area of rapeseed, as well as a minor increase in the cereal output and acreage. This means that, when the set-aside unit subsidy rate increases, farmers tend to intensify their production of these crops as they increase their set-aside area in parallel. This is confirmed by positive and significant elasticities of fertilizer and pesticide with respect to set-aside subsidy, equal to 0.0589 and 0.029, respectively. In the case of diester, an increase in the set-aside subsidy implies a reduction in both output and area of this crop, with a stronger substitution effect than the increase in cereal and rapeseed output and land use discussed above. Fertilizer demand increases with cereal and rapeseed prices (elasticities of 0.0117 and 0.0555, respectively) and decreases with the price of diester (elasticity of -0.1261). Fertilizer demand also increases with rapeseed subsidy (0.0350) and decreases with diester subsidy (elasticity of -0.1200) and does not vary significantly with the price of cereals (elasticity of -0.0010). Finally, demand for pesticide increases with diester price (elasticity of 0.0958) and decreases with the prices of cereals and rapeseed (elasticities of -0.0120 and -0.0548, respectively). Pesticide demand also increases with the subsidy for diester and decreases with the unit subsidy for cereals and rapeseed (elasticities of 0.0498 and -0.0316, respectively).

We compare our estimated elasticities with those found in the literature. Concerning the own-price elasticity of pesticide demand, the meta-analysis by [68] concludes that the median own-price pesticide demand elasticities in Europe is equal to -0.30. More specifically, in the case of France, [69] find an elasticity equal to -0.77 (-1.25 to -0.28) and [70] find a value of -0.17 (-0.24 to -0.10). Using a simulation procedure to evaluate the magnitude of tax change (or price increase) necessary to result in a given quantity of pesticide,

[71] find very low pesticide demand elasticity estimates, larger for specialized farms (between -0.026 and -0.049) than for diversified farms (between -0.011 and -0.023) in the short term. According to the literature review by [72], the own-price elasticity of pesticide is equal to -0.30 in the case of France. The value of our own price-elasticity of pesticide (-0.37) therefore falls within the values found in the literature for Europe and particularly for France. Our estimates show a lower value (-0.05) for the own-price elasticity of fertilizer demand than those in the literature for France: [70] (-0.50 to -0.16), [73] (-0.278) and [5] (-0.371). Moreover, our elasticity estimates for fertilizer and pesticide demand with respect to set-aside subsidy (0.0589 and 0.029, respectively) can also be compared with the empirical literature, which, however, remains limited on the subject. Estimates provided in the European context are fairly heterogeneous in magnitude, ranging from 0.12 for [5], 0.13 for [74], to 1.52 for [26].

Several conclusions emerge from our results on the environmental impact of a set-aside policy. All else being equal, a set-aside subsidy has a positive impact on farm-level fertilizer and pesticide demand: an increase in the set-aside subsidy of 1 percent implies an increase of 0.0589 percent in fertilizer demand and 0.029 percent in pesticide demand, respectively. However, the elasticity of chemical input demand with respect to the set-aside subsidy is calculated for total land with no distinction between cultivated and set-aside lands. The next subsection presents our second indicator that explicitly accounts for chemical input use on cultivated land only.

### 5.2 Chemical Input Intensity Elasticities with Respect to Set-aside subsidy

The second indicator we consider here is defined as the elasticity of chemical input quantity per unit of cultivated land with respect to the set-aside subsidy. This indicator accounts explicitly for land set-aside by considering chemical input use intensity per unit of cultivated land. It accounts for farmers' decisions about land set-aside following a change to the subsidy but is not dependent on crop distribution (only on total cultivated area). This indicator is calculated as follows:

$$\begin{aligned}
 \varepsilon_{II_k \bar{\tau}_s} &= \frac{\partial II_k}{\partial \bar{\tau}_s} \times \frac{\bar{\tau}_s}{II_k} = \frac{\partial \left( \frac{x_k}{L - l_{setaside}} \right)}{\partial \bar{\tau}_s} \times \frac{\bar{\tau}_s}{\left( \frac{x_k}{L - l_{setaside}} \right)} \\
 &= \frac{\partial x_k}{\partial \bar{\tau}_s} \times \frac{\bar{\tau}_s}{x_k} - \frac{\partial (L - l_{setaside})}{\partial \bar{\tau}_s} \times \frac{\bar{\tau}_s}{L - l_{setaside}} \\
 &= \frac{\partial x_k}{\partial \bar{\tau}_s} \times \frac{\bar{\tau}_s}{x_k} + \frac{\partial (l_{setaside})}{\partial \bar{\tau}_s} \times \frac{\bar{\tau}_s}{L - l_{setaside}} \\
 &= \varepsilon_{x_k \bar{\tau}_s} + \varepsilon_{l_{setaside} \bar{\tau}_s} * \frac{l_{setaside}}{L - l_{setaside}}. \tag{13}
 \end{aligned}$$



**Table 7** Elasticities of fertilizer and pesticide intensity with respect to unit set-aside subsidy

	$\epsilon_{II_f \bar{\tau}_s}$	$\epsilon_{II_p \bar{\tau}_s}$
PQML	0.0645*** (0.0018)	0.0346*** (0.0026)
SNPQML	0.0709** (0.0294)	0.0352 (0.0772)

$\epsilon_{l_{set-aside} \bar{\tau}_s}$  denotes the elasticity of set-aside area with respect to the set-aside subsidy.  $\epsilon_{II_f \bar{\tau}_s}$  and  $\epsilon_{II_p \bar{\tau}_s}$  denote, respectively, the elasticity of fertilizer and pesticide intensity with respect to the set-aside subsidy. PQML and SNPQML, respectively, denote parametric and semi-non-parametric quasi-maximum likelihood estimations. Robust standard errors are in parentheses (computed with the Delta method for input intensity elasticities)

\*, \*\* and \*\*\*, respectively, denote parameter significance at the 10, 5 and 1 percent level

The first term  $\epsilon_{x_k \bar{\tau}_s}$  is the demand elasticity of input  $k$  with respect to the set-aside subsidy and measures how a variation of 1 percent in the set-aside subsidy affects the percentage demand for input  $k$ . The second term  $\epsilon_{l_{set-aside} \bar{\tau}_s}$  is the elasticity of set-aside area with respect to the set-aside subsidy, which measures how a variation of 1 percent in the set-aside subsidy affects the percentage of the area set-aside. A positive  $\epsilon_{II_f \bar{\tau}_s}$  means that an increase of 1 percent in the set-aside subsidy will increase input demand  $k$  by unit of cropped land, implying that an intensification effect is observed.

Table 7 presents the results for input intensity elasticity of fertilizers  $\epsilon_{II_f \bar{\tau}_s}$  and pesticides  $\epsilon_{II_p \bar{\tau}_s}$  with respect to the set-aside subsidy for PQML and SNPQML specifications. These results show that input demand intensity elasticities with respect to a set-aside subsidy are positive and significant for both fertilizer (0.0645) and pesticide (0.0346) demand in the PQML model. This means that, when the set-aside subsidy increases, farmers tend to increase both their set-aside area  $l_{set-aside}$  and their input use (fertilizer and pesticide); to compensate for the loss due to a reduced crop area, farmers intensify their production by increasing their chemical input demand per hectare of crop area. In our case, this means that increasing the set-aside subsidy has an impact on agrochemical input intensification, which could have a negative impact on the environment (in terms, for example, of water contamination and biodiversity loss). Comparing the input elasticities for fertilizer and pesticide, the value of  $\epsilon_{II_f \bar{\tau}_s}$  is always higher than  $\epsilon_{II_p \bar{\tau}_s}$ . Fertilizer is usually considered a risk-increasing input, since it jointly increases the expected crop yield and its variance. In contrast, previous results in the literature regarding the direction of risk effects of pesticides are ambiguous. Möhring et al. [75] show that the indicator of pesticide choice affects the magnitude and sign of estimated risk effects.

### 5.3 Set-aside Policy Simulation

According to the new rules following the current CAP reform 2014–2020, farmers are required to implement greening measures or lose up to 30 percent of their basic payment scheme income. The greening rules cover three areas: crop diversification, ecological focus areas and non-intensification measures to maintain permanent grassland.

As the results of our elasticity estimation show, an increase in the set-aside subsidy could imply an increased demand for fertilizer and pesticide inputs. This means that a set-aside policy introduced as an EFA in order to preserve biodiversity could have some potential adverse environmental impacts due to intensification at the farm level. We use our elasticity estimates to simulate the impacts of a public policy that imposes a 5-percent increase in the set-aside area on demand for fertilizer and pesticide. To do this, from our elasticity of set-aside area with respect to its subsidy, we calculate the subsidy increase required to achieve a 5-percent increase in the set-aside area. This value is chosen with reference to the 2014–2020 CAP, which also contains such a requirement but with a major difference, the lack of area-based subsidy associated with land set-aside.

Let us start with the elasticity of set-aside area,  $\epsilon_{l_s \bar{\tau}_s}$ , with respect to set-aside subsidy,  $\tau_s$  on crop  $s$ :

$$\epsilon_{l_s \bar{\tau}_s} = \frac{\partial l_s}{\partial \bar{\tau}_s} \times \frac{\bar{\tau}_s}{l_s} = \frac{\partial l_s}{l_s} / \frac{\partial \bar{\tau}_s}{\bar{\tau}_s} \tag{14}$$

If we assume that  $\frac{\partial l_s}{l_s} = 0.05$  (5 percent), we can calculate the corresponding (equivalent) variation in the set-aside subsidy as

$$\frac{\partial \bar{\tau}_s}{\bar{\tau}_s} = \frac{0.05}{\epsilon_{l_s \bar{\tau}_s}}$$

We then use this variation of the set-aside subsidy above and the fertilizer and pesticide demand elasticities with respect to the set-aside subsidy ( $\epsilon_{f \bar{\tau}_s}$  and  $\epsilon_{p \bar{\tau}_s}$ , respectively) to calculate the corresponding fertilizer and pesticide demand variations, denoted  $\frac{\partial x_f}{x_f}$  and  $\frac{\partial x_p}{x_p}$ , respectively. We finally use these input demand variations and fertilizer and pesticide own-price elasticities ( $\epsilon_{x_f \bar{w}_f}$  and  $\epsilon_{x_p \bar{w}_p}$ , respectively) to calculate the corresponding fertilizer and pesticide “net” price variations ( $\frac{\partial w_f}{w_f}$  and  $\frac{\partial w_p}{w_p}$ ).

Strictly speaking, output and input market prices are exogenous (because farmers are price takers), but the final prices to the farmer are “net of tax”, that is, they incorporate the unit tax. As a result, price variations only concern the part of the final price associated with the policy instrument (the tax).

**Table 8** Simulation results of a 5-percent increase in set-aside area

	(1)	(2)	(3)	(4)	(5)
	$\frac{\partial \bar{\tau}_s}{\bar{\tau}_s}$	$\frac{\partial x_f}{x_f}$	$\frac{\partial x_p}{x_p}$	$\frac{\partial w_f}{w_f}$	$\frac{\partial w_p}{w_p}$
PQML	0.3381	0.0199	0.0098	0.3927	0.0265
		[0.0187;0.0211]	[0.0081;0.0115]	[0.3690;0.4163]	[0.0218;0.0312]

$\frac{\partial \bar{\tau}_s}{\bar{\tau}_s}$  denotes the variation of set-aside subsidy in percent.  $\frac{\partial x_f}{x_f}$  and  $\frac{\partial x_p}{x_p}$  denote, respectively, the fertilizer and pesticide demand variation in percent.  $\frac{\partial w_f}{w_f}$  and  $\frac{\partial w_p}{w_p}$ , respectively, denote the price of fertilizer and pesticide variation in percent. 95 percent confidence intervals obtained from standard errors of parameter estimates are presented in brackets. PQML denotes parametric quasi-maximum likelihood estimation

For this simulation, we use elasticities calculated from the parametric PQML model with fixed effects (Table 5). This choice is motivated by more efficient estimates (with respect to the semi-nonparametric model, as discussed above), with most elasticities significant at the 1 percent level. The results of this exercise are summarized in Table 8. Confidence intervals obtained from robust standard errors of parameter estimates are presented in brackets for the key expressions in this table.

We find that, in order to obtain a 5-percent increase in the set-aside area, we need to increase the set-aside subsidy rate by 33.81 percent. Using fertilizer and pesticide elasticities with respect to set-aside subsidy, our simulations show that an increase of the set-aside subsidy by 33.81 percent implies an increase in fertilizer demand ranging from 1.87 to 2.11 percent, and for pesticide, an increase in demand ranging from 0.81 to 1.15 percent (columns 3 and 4 of Table 8). This could have potentially adverse effects on the environment including, e.g., nitrogen runoff and ground water pollution. Using own price elasticities of fertilizer and pesticide demand, we can calculate the tax level necessary to offset such increase in fertilizer and pesticide demand. Our simulations show that such tax rates would range from 36.9 to 41.63 percent for fertilizer, and from 2.18 to 3.12 percent for pesticide (changes in input prices from the last two columns of Table 8). In line with most empirical papers dealing with elasticities of input use in agriculture (see, e.g., [68]), our results show that it requires a substantial tax level on pesticide and fertilizer to yield a significant reduction in input use: a tax on fertilizer between 36.9 and 41.63 percent to offset an increase in demand of between 1.87 to 2.11 percent and a tax on pesticide use between 2.18 and 3.12 percent to offset an increase in demand of between 0.81 and 1.15 percent. Note, however, that this result holds if a “homogeneous” value-added tax is applied. Average tax levels might be lower and the tax might be more efficient (in reducing potential environmental risks from pesticides) if tax levels are adjusted for heterogeneous pesticide properties, see [76] and [77]. These results are also in line with

those of [78], who considers a theoretical framework with a fiscal scheme consisting of both a tax on nitrogen application and a subsidy on land with cover crops. This paper shows that such a scheme is efficient in improving water quality and that it has the potential to balance the budget dedicated to the public policy.

## 6 Conclusion

Like any policy modifying marginal benefits of land, set-aside policies imply changes to crop choices and production practices, whose effects and intensity depend on various factors. These policies may even increase crop yield since farmers tend to use low-yield soils to meet set-aside requirements. As a consequence, average cultivated land quality may increase, implying as well an increase in aggregate crop yield per hectare. Such an effect may also be obtained because of input intensification with potentially adverse impacts on the environment, which would conflict with the initial objectives of the policy. To investigate these potential effects due to input intensification, this paper evaluates the effect of a set-aside policy, based on changes in agricultural practices and land use, for a sample of French farmers in the Meuse *Département* between 2006 and 2010.

We first derive fertilizer and pesticide input demand in the case of multiple crops, from a structural multi-output production model, estimated with both a parametric and a semi-nonparametric QML procedure, to account for multiple corner solutions. We use the most efficient estimator, the parametric QML, to compute elasticities with respect to the set-aside subsidy for two indicators: pesticide and fertilizer demand, and pesticide and fertilizer demand intensity per unit of cultivated land. We find that a policy targeting a 5-percent increase in the set-aside area results in an increase in fertilizer (resp. pesticide) demand from 1.87 to 2.11 (resp. 0.81 to 1.15) percent. Such policy would therefore be associated with chemical input intensification, with potentially adverse environmental effects in terms of biodiversity loss and water pollution. To offset such an intensive-margin

effect, taxes ranging from 36.9 to 41.63 (resp. 2.18 to 3.12) percent on fertilizer (resp. pesticide) would be necessary. Environmental effects are likely to be more harmful if the input demand intensity indicator increases for a reduced (cultivated) area. Note, however, that this indicator measures only the environmental pressure as a potential impact from input use, and not the actual environmental impact, which is likely to depend on a variety of factors (soil type, slope, climate, distance to surface or groundwater, and the characteristics of the local ecosystem.).

Because it is a nonpoint source pollution issue, only second-best outcomes may be achieved, using, e.g., indirect taxation (pesticide and fertilizer sales, output level, land use, etc.). Policies consisting of taxing chemical inputs, alongside subsidizing land set-aside, are often advocated as feasible second-best policies, because the first-best policy of taxing environmental damage would be prohibitively costly to implement (monitoring and management costs to the environmental regulator, valuing environmental damages, etc.). Taxing fertilizer and pesticide use to correct for market failures (externalities including water contamination and human poisoning) is generally considered a cost-effective policy in theory (as opposed to command-and-control policies), provided implementation costs are limited and the tax level does not deviate excessively from the optimal (first-best) level. Note also that reducing pesticide use through a tax scheme may be beneficial to the farmer in the long run, because of a potential reduction in, e.g., resistance of plant pests. Although relevant in principle, accommodating such extension would require additional information at the plot or the farm level on the benefits of such tax in terms of pest resistance reduction, in a way similar to our discussion above on a differentiated, risk-specific pesticide tax (see the end of Sect. 5.3).

Revenue streams generated by fertilizer and pesticide taxes can be earmarked to subsidize more sustainable agricultural practices [79]. This implies that in principle, considering a policy consisting in complementing an input tax with a subsidy on set-aside would correspond to the earmarking strategy above. However, the objective of the tax simulation considered here is to offset the increase in input demand following intensification, to illustrate the magnitude of the tax on fertilizer and pesticide as equivalent policies. In other words, the level of the input tax required to offset the negative consequences of a set-aside policy can also be interpreted as the tax level that would be necessary to reduce fertilizer and pesticide use by the same amount (as the increase in demand in the first place). We do not discuss the relative advantages of tax vs. area-based subsidies as policy instruments, which can be found, e.g., in [80], with a discussion on alternative policy instruments that can be considered to correct for market failures associated with agricultural pollution.

Our analysis could be extended in several directions. First, the model could be improved by incorporating other policy instruments, for example, in the case of the recent European Common Agricultural Policy reform, the number of crops in rotation and the proportion of grassland area. Another extension would involve linking our production model to observed farm-level environmental variables, such as water quality and biodiversity. Second, crop rotations may be considered in an extended framework where previous crop decisions could be accounted for in current production decisions. As crop rotations may impose restrictions on future farmer choices in subsequent years, such extension may provide interesting insights as to the potential impact of policies designed to modify agricultural land use decisions. A further difficulty however is that plot-level data are necessary, to fully capture the benefits associated with crop rotations on a set of land plots. Third, concerning our indicators capturing environmental pressure from cropping practices and land use, we could consider other pesticide use indicators [81] or the application of different weights to each crop when computing these indicators. Based on “technical” parameters, the extended indicators could then better reflect heterogeneous environmental risks associated with each crop.

From an econometric viewpoint, several extensions could also be considered, starting with the correction of possible errors-in-variables (EIV) bias when deflating the price system by the price of seed as numeraire, as proposed by [82]. Furthermore, to obtain results that do not vary with the choice of numeraire, an extension would be to consider a symmetric normalized quadratic profit function, as in [83], however with an additional computational cost.<sup>18</sup> This is left for future research.

**Acknowledgements** We thank the Editor of the journal and two anonymous referees for their very detailed and helpful comments. We also thank the Centre de Gestion et d’Economie Rurale de la Meuse/CER for providing the data.

**Author Contributions** All authors contributed equally to the paper.

**Funding** Funding from the Agence Nationale de la Recherche as part of the “Investments d’Avenir Program” within STIMUL project (ANR-11-LABX-0034), Cland Institut de convergence (ANR-16-CONV-0003) and Fast projet (ANR-16-IDEX-0006) are gratefully acknowledged. Funding from ANR (Agence Nationale de la Recherche, France) under grant Investissements d’Avenir program (ANR-17-EURE-0010) is also gratefully acknowledged

**Data Availability** We used individual data for farmers from Meuse Département, made available to us thanks to an agreement with Centre d’Economie Rurale de la Meuse. These data are subject to statistical confidentiality and therefore cannot be disseminated.

<sup>18</sup> Paudel and McIntosh [89] show that estimation results do not always vary significantly with the choice of the numeraire.

**Code Availability** Code is available and could be shared. We used Fortran F98 (Absoft compiler for Windows with IMSL statistical and mathematical libraries – procedures SPLEZ and QD2VL for spline evaluation).

## Declarations

**Conflicts of Interest** No conflicts of interest.

## References

- Chen, X. (2007). Large sample sieve estimation of semi-nonparametric models. In Heckman, J., & Leamer, E. (Eds.) *Handbook of Econometrics* (vol. 6, ch. 76). Elsevier, London.
- Yen, S. T., Lin, B.-H., & Smallwood, D. M. (2003). Quasi-and simulated-likelihood approaches to censored demand systems: food consumption by food stamp recipients in the United States. *American Journal of Agricultural Economics*, 85(2), 458–478.
- Fezzi, C., & Bateman, I. (2011). Structural agricultural land use modeling for spatial agro-environmental policy analysis. *American Journal of Agricultural Economics*, 93(4), 1168–1188.
- Gouriéroux, C., & Monfort, A. (1993). Large sample sieve estimation of semi-nonparametric models. In Maddala, C. R. R. G. S., & Vinod, H. D. (Eds.), *Handbook of Statistics* (vol. 11, ch. 12). Elsevier, London.
- Lacroix, A., & Thomas, A. (2011). Estimating the environmental impact of land and production decisions with multivariate selection rules and panel data. *American Journal of Agricultural Economics*, 93(3), 784–802.
- Montoya-Blandon, S., & Jacho-Chavez, D. (2020). Semiparametric quasi maximum likelihood estimation of the fractional response model. *Economics Letters*, 186, 108769.
- Van Buskirk, J., & Willi, Y. (2004). Enhancement of farmland biodiversity within set-aside land. *Conservation Biology*, 18(4), 987–994.
- Kleijn, D., Berendse, F., Smit, R., & Gilissen, N. (2001). Agri-environment schemes do not effectively protect biodiversity in Dutch agricultural landscapes. *Nature*, 413(6857), 723–725.
- European Court of Auditors. (2017). Greening: a more complex income support scheme, not yet environmentally effective. *Special Report No 21/2017. European Court of Auditors, Luxembourg*.
- Rozakis, S., & Sourie, J.-C. (2005). Micro-economic modelling of biofuel system in France to determine tax exemption policy under uncertainty. *Energy Policy*, 33(2), 171–182.
- European Commission. (2013). Overview of CAP reform 2014–2020. *Agricultural Policy Perspectives Brief*, 5.
- Matthews, A. (2013). Greening agricultural payments in the EU's common agricultural policy. *Bio-based and Applied Economics*, 2(1), 1–27.
- Hauck, J., Schleyer, C., Winkler, K. J., & Maes, J. (2014). Shades of greening: reviewing the impact of the new EU agricultural policy on ecosystem services. *Change and Adaptation in Socio-Ecological Systems*, 1, 1.
- Meredith, S., & Kolenda, E. (2021). CAP dialogue negotiations briefing. Tech. rep.
- Fraser, I., & Stevens, C. (2008). Nitrogen deposition and loss of biological diversity: Agricultural land retirement as a policy response. *Land Use Policy*, 25(4), 455–463.
- Love, H., & Foster, W. (1990). Commodity program slippage rates. *Western Journal of Agricultural Economics*, 15(2), 272–281.
- Ball, V. E., Bureau, J.-C., Eakin, K., & Somwaru, A. (1997). CAP reform: modelling supply response subject to the land set-aside. *Agricultural Economics*, 17(2–3), 277–288.
- Choi, J.-S., & Helmberger, P. G. (1993). How sensitive are crop yields to price changes and farm programs? *Journal of Agricultural and Applied Economics*, 25, 237–244.
- Fraser, I., & Waschik, R. (2005). Agricultural land retirement and slippage: lessons from an Australian case study. *Land Economics*, 81(2), 206–226.
- Rygnestad, H., & Fraser, R. (1996). Land heterogeneity and the effectiveness of CAP set-aside. *Journal of Agricultural Economics*, 47(1–4), 255–260.
- Vannini, L., Gentile, E., Bruni, M., Loi, A., Aragrande, M., Theuvsen, L., Haman, K., Venturi, G., Monti, A., Canali, M., & Bernini, C. (2008). Evaluation of the set aside measure 2000 to 2006: Final report. Tech. rep.
- Wu, J. (2000). Slippage effects of the Conservation Reserve Program. *American Journal of Agricultural Economics*, 82(4), 979–992.
- Gordard, R. (2013). Profit-maximizing land-use revisited: The testable implications of non-joint crop production under land constraint. *American Journal of Agricultural Economics*, 95, 1109–1121.
- Reicheldefer, K., & Bogges, W. G. (1988). Government decision making and program performance: The case of the conservation reserve program. *American Journal of Agricultural Economics*, 70(1), 1–11. <https://doi.org/10.2307/1241970>
- Hendricks, N. P., & Er, E. (2018). Changes in cropland area in the United States and the role of CRP. *Food Policy*, 75, 15–23.
- Laukkanen, M., & Nauges, C. (2014). Evaluating greening farm policies: A structural model for assessing agri-environmental subsidies. *Land Economics*, 90(3), 458–481.
- Salles, J., Teillard, F., Tichit, M., & Zanella, M. (2017). Land sparing versus land sharing: an economist's perspective. *Regional Environmental Change*, 17(5), 1455–1465.
- Guyomard, H., Baudry, M., & Carpentier, A. (1996). Estimating crop supply response in the presence of farm programmes: Application to the cap. *European Review of Agricultural Economics*, 23(4), 401–420.
- Moro, D., & Sckokai, P. (1999). Modelling the CAP arable crop regime in Italy: Degree of decoupling and impact of agenda 2000. *Cahiers d'Economie et Sociologie Rurales*, 0(53), 49–73.
- Tobin, J. (1958). Estimation of relationships for limited dependent variables. *Econometrica*, 26(1), 24–36.
- Pudney, S. (1989). *Modelling individual choice: The econometrics of corners, kinks and holes*. Oxford: Basil Blackwell.
- Heckman, J. J. (1978). Dummy endogenous variables in a simultaneous equation system. *Econometrica*, 46(6), 931–959.
- Wales, T. J., & Woodland, A. D. (1983). Estimation of consumer demand systems with binding non-negativity constraints. *Journal of Econometrics*, 21(3), 263–285.
- Lee, L.-F., & Pitt, M. M. (1986). Microeconomic demand system with binding nonnegativity constraints: The dual approach. *Econometrica*, 54(5), 1237–1242.
- Lee, L., & Pitt, M. (1987). Microeconomic models of rationing, imperfect markets, and non-negativity constraints. *Journal of Econometrics*, 36, 89–110.
- Arndt, C. (1999). Demand for herbicide in corn: An entropy approach using micro-level data. *Journal of Agricultural and Resource Economics*, 24, 204–21.
- Golan, A., Perloff, J., & Shen, E. (2001). Estimating a demand system with nonnegativity constraints: Mexican meat demand. *The Review of Economics and Statistics*, 83(3), 541–550.
- Perali, F., & Chavas, J.-P. (2000). Estimation of censored demand equations from large cross-section data. *American Journal of Agricultural Economics*, 82(4), 1022–1037.

39. Meyerhoefer, C. D., Ranney, C. K., & Sahn, D. E. (2005). Consistent estimation of censored demand systems using panel data. *American Journal of Agricultural Economics*, 87(3), 660–672.
40. Shonkwiler, J., & Yen, S. (1999). Two-step estimation of a censored system of equations. *American Journal of Agricultural Economics*, 81(4), 972–82.
41. Yen, S. T., & Lin, B.-H. (2006). A sample selection approach to censored demand systems. *American Journal of Agricultural Economics*, 88(3), 742–749.
42. Kao, C., Lee, L.-F., & Pitt, M. (2001). Simulated maximum likelihood estimation of linear expenditure system with binding non-negativity constraints. *Annals of Economics and Finance*, 2, 215–235.
43. Chakir, R., & Thomas, A. (2003). Simulated maximum likelihood estimation of firm demand systems with corner solutions and panel data. *Revue d'Economie Politique*, 6, 773–997.
44. Millimet, D. L., & Tchernis, R. (2008). Estimating high-dimensional demand systems in the presence of many binding non-negativity constraints. *Journal of Econometrics*, 147(2), 384–395.
45. Platoni, S., Sckokai, P., & Moro, D. (2012b). Panel data estimation techniques and farm-level data models. *American Journal of Agricultural Economics*, 1202–1217.
46. Platoni, S., Sckokai, P., & Moro, D. (2012a). A note on two-way ECM estimation of SUR systems on unbalanced panel data. *Econometric Reviews*, 31(2), 119–141.
47. Chambers, R. G., & Just, R. E. (1989). Estimating multioutput technologies. *American Journal of Agricultural Economics*, 71(4), 980–995.
48. Thomas, A. (2003). A dynamic model of on-farm integrated nitrogen management. *European Review of Agricultural Economics*, 30(4), 439–460.
49. Hennessy, D. A. (2006). On monoculture and the structure of crop rotations. *American Journal of Agricultural Economics*, 88, 900–914.
50. Wooldridge, J. M. (1995). Selection corrections for panel data models under conditional mean independence assumptions. *Journal of Econometrics*, 68(1), 115–132.
51. Papke, L. E., & Wooldridge, J. M. (1996). Econometric methods for fractional response variables with an application to 401 (k) plan participation rates. *Journal of Applied Econometrics*, 11(6), 619–632.
52. Duncan, G. (1986). A semi-parametric censored regression estimator. *Journal of Econometrics*, 32, 5–34.
53. Gallant, A., & Nychka, D. (1987). Semi-nonparametric maximum likelihood estimation. *Econometrica*, 55, 363–390.
54. Schwiebert, J. (2013). Sieve maximum likelihood estimation of a copula-based sample selection model. University of Hannover. Working paper.
55. Chen, X., Fan, Y., & Tsyrennikov, V. (2006). Efficient estimation of semiparametric multivariate copula models. *Journal of the American Statistical Association*, 101, 1228–1240.
56. Chamberlain, G. (1982). Multivariate regression models for panel data. *Journal of Econometrics*, 18(1), 5–46.
57. Agreste. (2017). Synthèse départementale, la meuse (55). ministry of food, agriculture and forestry: Châlons-en-champagne. Tech. rep., Agreste-Grand-Est.
58. Britz, W., Heckeles, T., & Pérez, I. (2006). Effects of decoupling on land use: an EU wide, regionally differentiated analysis. *Agrarwirtschaft*, 55(5), 215.
59. Moro, D., & Sckokai, P. (2013). The impact of decoupled payments on farm choices: Conceptual and methodological challenges. *Food Policy*, 41, 28–38.
60. Boulanger, P. (2010). Distribution of agricultural support: selected French evidences. In *OECD Workshop on the Disaggregated Impacts of CAP Reform*, March (pp. 10–11).
61. Born, B., & Breitung, J. (2016). Testing for serial correlation in fixed-effects panel data models. *Econometric Reviews*, 35, 1290–1316.
62. Vuong, Q. H. (1989). Likelihood ratio tests for model selection and non-nested hypotheses. *Econometrica*, 57(2), 307–33.
63. Lavergne, P., & Thomas, A. (2005). Semiparametric estimation and testing in a model of environmental regulation with adverse selection. *Empirical Economics*, 30(1), 171–192.
64. Shortle, J. S., & Horan, R. D. (2001). The economics of nonpoint pollution control. *Journal of Economic Surveys*, 15(3), 255–289.
65. Green, R. E., Cornell, S. J., Scharlemann, J. P., & Balmford, A. (2005). Farming and the fate of wild nature. *Science*, 307(5709), 550–555.
66. Didham, R. K., Barker, G. M., Bartlam, S., Deakin, E. L., Denmead, L. H., Fisk, L. M., et al. (2015). Agricultural intensification exacerbates spillover effects on soil biogeochemistry in adjacent forest remnants. *PLoS One*, 10, 1.
67. D'Amario, S. C., Rearick, D. C., Fasching, C., Kembel, S. W., Porter-Goff, E., Spooner, D. E., et al. (2019). The prevalence of nonlinearity and detection of ecological breakpoints across a land use gradient in streams. *Scientific reports*, 9(1), 1–11.
68. Böcker, T. G., & Finger, R. (2017). A meta-analysis on the elasticity of demand for pesticides. *Journal of Agricultural Economics*, 68(2), 518–533.
69. Jacquet, F., Butault, J.-P., & Guichard, L. (2011). An economic analysis of the possibility of reducing pesticides in french field crops. *Ecological economics*, 70(9), 1638–1648.
70. Femenia, F., & Letort, E. (2016). How to significantly reduce pesticide use: An empirical evaluation of the impacts of pesticide taxation associated with a change in cropping practice. *Ecological Economics*, 125, 27–37.
71. Ayoub, K., & Vigeant, S. (2020). Can we really use prices to control pesticide use? Results from a nonparametric model. *Environmental Modeling & Assessment*, 1–16.
72. Skevas, T., Lansink, O., & Stefanou, S. (2013). Designing the emerging EU pesticide policy: A literature review. *NJAS-Wageningen Journal of Life Sciences*, 64, 95–103.
73. Bayramoglu, B., & Chakir, R. (2016). The impact of high crop prices on the use of agro-chemical inputs in France: A structural econometric analysis. *Land Use Policy*, 55, 204–211.
74. Simola, A. (2019). *Essays on agricultural policies and land use*. PhD thesis, University of Helsinki, Finland.
75. Möhring, N., Bozzola, M., Hirsch, S., & Finger, R. (2020). Are pesticides risk decreasing? The relevance of pesticide indicator choice in empirical analysis. *Agricultural Economics*, 51(3), 429–444.
76. Kudsk, P., Jorgensen, L. N., & Orum, J. E. (2018). Pesticide load—a new Danish pesticide risk indicator with multiple applications. *Land Use Policy*, 70, 384–393.
77. Finger, R., Möhring, N., Dalhaus, T., & Böcker, T. (2017). Revisiting pesticide taxation schemes. *Ecological Economics*, 134, 263–266.
78. Martin, E. (2019). Cover crops and water quality. *Environmental Modeling & Assessment*, 24(6), 605–623.
79. Weitzman, M. L., & Spence, A. M. (1978). *Regulatory Strategies for Pollution Control*. Cambridge, MA, London, UK: M.I.T. Press.
80. Sterner, T. (2003). *Policy instruments for environmental and natural resource management*. Washington, DC: Resources for the Future.
81. Möhring, N., Gaba, S., & Finger, R. (2019). Quantity based indicators fail to identify extreme pesticide risks. *Science of the Total Environment*, 646, 503–523.
82. Pope, R., LaFrance, J., & Just, R. (2007). Imperfect price deflation in production systems. *American Journal of Agricultural Economics*, 89, 738–754.

83. Kohli, U. (1993). A symmetric normalized quadratic GNP function and the US demand for imports and supply of exports. *International Economic Review*, *34*, 243–255.
84. Hansen, L. (2007). Conservation reserve program: Environmental benefits update. *Agricultural and Resource Economics Review* *36*(2), 267–280. <https://doi.org/10.1017/S1068280500007085>
85. Diewert, W., & Wales, T. (1987). Flexible functional forms and global curvature conditions. *Econometrica*, *55*(1), 43–68.
86. Sckokai, P., & Moro, D. (2009). Modelling the impact of the CAP single farm payment on farm investment and output. *European Review of Agricultural Economics*, *36*(3), 395–423.
87. Lansink, A. O., & Carpentier, A. (2001). Damage control productivity: An input damage abatement approach. *Journal of Agricultural Economics*, *52*(3), 11–22.
88. Dessart, F. J., Barreiro-Hurlé, J., & van Bavel, R. (2019). Behavioural factors affecting the adoption of sustainable farming practices: a policy-oriented review. *European Review of Agricultural Economics*, *46*(3), 417–471.
89. Paudel, K. P., & McIntosh, C. S. (2007). Numeraire choice in agricultural supply analysis. *Applied Economics*, *37*, 1209–1214.

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.