Estimating the Environmental Impact of Land and Production Decisions with Multivariate Selection Rules and Panel Data

Anne Lacroix and Alban Thomas

We propose an estimation procedure dealing with multivariate selection and dynamics in land choice and unobserved heterogeneity in structural and crop selection equations. The method is applied to a multi-output profit function with multiple corner solutions and farm-level panel data. Estimation results reveal that the single selection rule associated with the Tobit procedure is rejected in favor of our multivariate selection treatment. Inconsistent estimates are expected with the Tobit procedure, or if farmer fixed effects are omitted, which can have important consequences for policy recommendations, as illustrated in an evaluation of nitrogen runoff elasticities with respect to economic variables.

Key words: environmental impact, land use, panel data, sample selection.

JEL codes: C33, C34, Q12, Q15.

Evaluating the impact of price and policy changes on land use and farmers' output is a major issue in applied agricultural economics. Consistent predictions of production choices are necessary, in particular, for environmental agencies to design more efficient policies and to properly evaluate environmental outcomes of agricultural policy reforms. Agricultural supply models generally assume that farmers maximize total profit over a set of possible crops, given predetermined prices and public compensatory payments. Farmers' decisions are, however, also conditioned by agronomic constraints regarding soil fertility and pest control, involving in particular previous land-use choices as in crop rotations. For example, Wu et al. (2004) find that crop choices are relatively inelastic to prices in a multinomial logit model of crop and tillage-practice decisions and explain this result, in part, by "agronomic (rotational) constraints" (p. 31). Predictions of farmers' decisions are likely to be inconsistent if changes in cropping systems and the influence of previous land for crops are not accounted for. For example, the own-price elasticity of output may be overestimated by a single-crop model because substitute crops are not considered. Similarly, the environmental impact of corn production may be underestimated by a static economic model of agricultural supply, if nitrogen carryover from previously planted crops is high.

Using repeated cross-sections (panel data) for modelling agricultural production at the individual level allows the analyst to partly compensate for the lack of soil and climate information. Panel data analysis can accommodate dynamic patterns in land for crops (which include pest control and soil fertility objectives), and it facilitates the control of unobserved individual heterogeneity. The latter is often a significant random component of land-use and agricultural production models capturing, e.g., soil, climate, and farmer-specific determinants of crop yield and profitability which are not explained by price and other economic variables (Wu et al. 2004). Adapting multi-output models of agricultural supply to panel data is therefore an interesting approach when the objective is to account (at least partly) for agronomic constraints which affect production decisions.

Two econometric issues seem particularly critical, however, when dealing with farm-level panel data: (*a*) corner solutions are frequent with individual data in a multiple crop setting;

Anne Lacroix is an agricultural economist at GAEL, INRA Grenoble. Alban Thomas is senior researcher at Toulouse School of Economics, INRA-LERNA Toulouse. We thank Alain Carpentier, Hervé Guyomard, Thierry Magnac, Stéphane Straub, and three anonymous referees for useful comments.

Amer. J. Agr. Econ. 93(3): 784–802; doi: 10.1093/ajae/aar008
Received May 2010; accepted January 2011; published online March 7, 2011

and (*b*) unobserved individual heterogeneity may be correlated with the model's explanatory variables. The first issue leads to sample selection bias if unobserved random terms underlying the choice of crops and production decisions are correlated. In a multi-crop setting, selection can be multivariate if the probability of crop choice and output variables are correlated across crops. The second issue implies that farmer-specific individual effects need to be controlled for, to avoid estimation bias and therefore incorrect inference.

To our knowledge, no prior work has considered the full implication of unobserved individual effects for multivariate selection of crops, in the case of agricultural production and land choice. The objective of the article is to address the following question: to what extent is overlooking the dynamic and multivariate nature of crop selection and unobserved individual effects in estimating multi-output production models affecting the inference on land use and production decisions? We propose a procedure for estimating a multi-crop production model with multivariate selection of crops, allowing for unobserved farmer individual effects in output, land and selection equations, and a dynamic pattern in land-use decisions. Issues of selection and unobserved random terms can be combined in practice, as illustrated in the following example.

Consider a farmer engaged in multi-crop production according to a particular cropping system (e.g., cereals, oilseed, protein crops). The farmer will decide to grow a particular crop, say, corn, depending on its relative profitability, on crops planted previously to account for agronomic constraints (nitrogen management, pest control) and also on unobserved determinants (to the analyst). When this decision is taken, the farmer will target the profit-maximizing output level depending on economic variables but also on unobserved determinants of crop yield. In both decisions, unobserved factors can be either farmer specific (soil quality, drainage of land plots, technical familiarity with the crop, etc.) or period dependent (climate, pest). If farmers are price takers, period-dependent random terms are not likely to influence price or policy variables. On the other hand, idiosyncratic farmers' characteristics may be correlated with the permanent level of such variables. For example, compensatory payments per unit of land planted with corn may be indexed on average regional corn yield, as in the European Common Agricultural Policy (CAP), or corn output prices may be systematically lower in regions where agronomic conditions are more favorable to this crop. Sample selection will occur if unobserved determinants of the decision to grow corn are correlated with the unobserved components of corn output. For example, soil quality can facilitate the introduction of corn into the cropping system, while ensuring a higher average crop yield. Moreover, a climatic shock (e.g., poor rainfall) may influence corn yield while lowering both the probability that corn and oilseed are planted. This is a case of multivariate selection of crops, as unobserved components may affect the choice of several crops by the farmer while influencing the final outcome in terms of land for crop and crop output level.

This article makes several contributions to the literature on applied agricultural production. First, we control for the influence of crop rotations on land-use and output decisions at the farm level by augmenting the production model with a crop-selection mechanism that depends on past proportions of crops in total land. Crop rotations have often been neglected in agricultural production models. Exceptions are the studies by Hennessy (2006), who provides a formal analysis of crop rotations in relation with carryover effects, time rationing, and risk aversion, and Thomas (2003) who estimates a structural model to identify simple crop rotations. Crop rotations are motivated on agronomic grounds for weed and pest control and soil fertility management (see Hennessy 2006), and may be considered the main reason behind selection issues in multi-crop contexts. Another reason, however, is simply the fact that some crops may temporarily be less attractive for a farmer, due to a change in relative output prices (or relative crop-specific costs).

Second, we propose a consistent method for dealing with multivariate selection of crops, by accounting for the possible correlation between unobserved components in crop selection and structural (production) equations. Very few applications in production analysis deal with sample selection in a multi-crop setting, and they generally control for selection bias on a univariate, crop-specific basis. For example, Moore and Negri (1992) deal with corner solutions in acreage and output decisions by estimating crop-specific equations by Tobit because "estimating Tobit regressions using a system framework is computationally intractable" (p. 34). Moro and Sckokai (1999) estimate a normalized quadratic multi-output profit function with a type 2 Tobit procedure. Note that this univariate correction procedure is more widely used in applied demand than in production analysis (Shonkwiler and Yen 1999; Yen and Lin 2005). This approach would be relevant under the assumption that expected land acreage or output of a crop selected by a farmer do not depend on the selection process underlying the choice of other crops. However, agronomic and economic factors may include unobserved components common to several crops, both in selection and production equations. In that case, nonzero correlations would exist between land and output equations for a particular crop on the one hand, and selection equations for other crops on the other. The type 2 Tobit approach above will then produce biased parameter estimates if omitted correction terms are correlated with explanatory variables, and be inefficient because it would not use the full information about the error correlation (see Yen 2005). Our method includes the crop-specific type 2 Tobit procedure as a special case, allowing us to perform a specification test of sample selection of crops.

Third, the estimation method is easier to implement than most approaches to multivariate selection found in the literature. The first approach is to estimate a structural model using inequality conditions between market and reservation prices to represent selection, along the lines of Lee and Pitt (1987). It has been applied mostly to consumer demand analysis and the agro-food industry (e.g., Chakir and Thomas 2003; Millimet and Tchernis 2009; Tiffin and Arnoult 2008; Yen 2005) and requires computer-intensive numerical methods. A second approach is to decompose the problem into a series of bivariate Tobits. Perali and Chavas (2000) propose a heteroskedastic Tobit model with a jackknife procedure, and Yen, Lin, and Smallwood (2003) use quasi-maximum likelihood to deal with multivariate selection, in the case of food consumption and in a cross-section context. These approaches are not robust, however, to heteroskedasticity of unknown form, and are still computer intensive. Our procedure to control for multivariate selection uses standard panel-data Probit and system of equations estimation procedures, and completely avoids computer-intensive (e.g., simulation-based estimation) approaches. It accommodates heteroskedasticity in the selection equation Probits, and a robust variance-covariance matrix of production parameter estimates is easily computed. Farmers' individual effects are taken care of in selection and structural equations by a fixed-effect procedure along the lines of Wooldridge (1995). Our estimation method is based on the assumption of no-correlation between crop-specific selection equations, given observed determinants of land use. This is a practical limitation of the method, but it is the price to pay for the simplification of the econometric procedure we propose. Nevertheless, such assumption is easily testable and should not be so difficult to satisfy in practice, as our empirical application demonstrates.

As an empirical application, we estimate a normalized quadratic profit function with multivariate crop selection and panel data, with output and input prices, total land and compensatory payments (subsidies) as explanatory variables. The multivariate model of crop selection and the system of structural equations is estimated on a panel of French farmers for the years 1995–2001. To examine the consequences of model misspecification (omitting individual effects or dynamics in land for crops, assuming a single crop selection rule), we compare our estimates with alternative specifications obtained as special cases. Many agricultural production models address estimation of land and output models and their implications for the environment by using aggregate data (Guyomard et al. 1996; Langpap, Hascic, and Wu 2008; Miller and Plantinga 1999; Orazem and Miranowski 1994; Plantinga 1996). As discussed in Just (1993) however, individual-level data allow for a much better understanding of producer decisions than aggregate data. We therefore also compare production elasticities obtained from estimation of the same model, but with aggregate data instead of farm-level data. Finally, in order to assess the consequence of model misspecification on policy recommendations, we estimate nitrogen-runoff elasticities from an environmental simulator, and compare the results of our method with alternative model specifications.

The Multi-Output Profit Function for Land Use and Input Choice

Consider a price-taking farmer with total land denoted L. The farmer allocates L to a set of C possible crops, each entailing a positive subsidy payment proportional to land area. Let c, $c = 1, 2, \ldots C$ denote the crop index, l_c denotes land used for crop c, p_c and τ_c denote output production price and unit subsidy rate (per hectare) respectively, and q_c denotes crop yield (per hectare). We denote by W_k and r_k, $k = 1, 2, \ldots, K$ respectively the kth component of the input vector (farm-level total quantity of input k) and the corresponding unit input price. The profit function is

$$(1) \qquad \Pi = \sum_{c=1}^{C} l_c[p_c q_c + \tau_c] - \sum_{k=1}^{K} r_k W_k$$

whose maximization under the constraint on total land $\sum_{c=1}^{C} l_c = L$ yields optimal land, output and input decisions $l_c(\boldsymbol{p}, \boldsymbol{\tau}, \boldsymbol{r}, L)$, $q_c(\boldsymbol{p}, \boldsymbol{\tau}, \boldsymbol{r}, L)$ and $W_k(\boldsymbol{p}, \boldsymbol{\tau}, \boldsymbol{r}, L)$. While a primal approach implies specifying crop yield functions, the dual approach, on the contrary, is based on the specification of a flexible form for profit, $\Pi(\boldsymbol{p}, \boldsymbol{\tau}, \boldsymbol{r}, L)$. Any well-behaved profit function must satisfy the following regularity conditions: homogeneity of degree one in prices, convexity in prices, monotonicity and symmetry. Because our profit function also depends on land-related payment rates, we need to impose the additional land adding-up condition:

$$(2) \qquad \sum_{c=1}^{C} l_c = L \Leftrightarrow \sum_{c=1}^{C} \frac{\partial l_c}{\partial p_{c'}} = \sum_{c=1}^{C} \frac{\partial l_c}{\partial \tau_{c'}} = \sum_{c=1}^{C} \frac{\partial l_c}{\partial r_k} = 0 \quad \forall c', \ \forall k \quad \text{and} \quad \sum_{c=1}^{C} \frac{\partial l_c}{\partial L} = 1.$$

Condition (2) as well as homogeneity of profit with respect to prices is easily imposed when a normalised quadratic form for profit is used. Also, negative profit values are easily dealt with, and profit is complemented with a system of supply and demand equations requiring knowledge of output and input quantities, in addition to price information.

The quadratic profit function is

$$(3) \qquad \bar{\Pi} = \beta_0 + \sum_{c=1}^{C} \beta_c^p \bar{p}_c + \sum_{c=1}^{C} \beta_c^\tau \bar{\tau}_c + \sum_{k=1}^{K-1} \beta_k^r \bar{r}_k + \sum_{c=1}^{C} \sum_{k=1}^{K-1} \beta_{ck}^{pr} \bar{p}_c \bar{r}_k + \sum_{c=1}^{C} \sum_{k=1}^{K-1} \beta_{ck}^{\tau r} \bar{\tau}_c \bar{r}_k + \frac{1}{2} \sum_{c=1}^{C} \sum_{c'=1}^{C} \beta_{cc'}^{pp} \bar{p}_c \bar{p}_{c'} + \frac{1}{2} \sum_{c=1}^{C} \sum_{c'=1}^{C} \beta_{cc'}^{\tau\tau} \bar{\tau}_c \bar{\tau}_{c'} + \sum_{c=1}^{C} \sum_{c'=1}^{C} \beta_{cc'}^{p\tau} \bar{p}_c \bar{\tau}_{c'} + \frac{1}{2} \sum_{k=1}^{K-1} \sum_{k'=1}^{K-1} \beta_{kk'}^{rr} \bar{r}_k \bar{r}_{k'} + \sum_{c=1}^{C} \gamma_c^p \bar{p}_c L + \sum_{c=1}^{C} \gamma_c^\tau \bar{\tau}_c L + \sum_{k=1}^{K-1} \gamma_k^r \bar{r}_k L,$$

where the upper bar indicates normalized profit and price variables, i.e., $\bar{\Pi} = \Pi / r_K, \bar{p}_c = p_c / r_K, \bar{\tau}_c = \tau_c / r_K, \bar{r}_k = r_k / r_K$, and r_K is the price of the *numeraire*.

Differentiating profit with respect to prices and unit subsidy rates yields

$$(4) \qquad l_c q_c = \frac{\partial \Pi}{\partial \bar{p}_c} = \beta_c^p + \sum_{k=1}^{K-1} \beta_{ck}^{pr} \bar{r}_k + \sum_{c'=1}^{C} \beta_{cc'}^{pp} \bar{p}_{c'} + \sum_{c'=1}^{C} \beta_{cc'}^{p\tau} \bar{\tau}_{c'} + \gamma_c^p L, \quad c = 1, \ldots, C,$$

$$(5) \qquad l_c = \frac{\partial \Pi}{\partial \bar{\tau}_c} = \beta_c^\tau + \sum_{k=1}^{K-1} \beta_{ck}^{\tau r} \bar{r}_k + \sum_{c'=1}^{C} \beta_{cc'}^{p\tau} \bar{p}_{c'} + \sum_{c'=1}^{C} \beta_{cc'}^{\tau\tau} \bar{\tau}_{c'} + \gamma_c^\tau L, \quad c = 1, \ldots, C,$$

$$(6) \qquad -W_k = \frac{\partial \Pi}{\partial \bar{r}_k} = \beta_k^r + \sum_{c=1}^{C} \beta_{ck}^{pr} \bar{p}_c + \sum_{c=1}^{C} \beta_{ck}^{\tau r} \bar{\tau}_c + \sum_{k'=1}^{K-1} \beta_{kk'}^{rr} \bar{r}_{k'} + \gamma_k^r L, \quad k = 1, \ldots, K-1,$$

from which elasticities of land, output and input with respect any price (or subsidy rate), can be computed by multiplying the corresponding parameter (price or subsidy rate in the land or output equation) by the ratio of the (normalized) price over land or output level.

From the property of linear homogeneity in prices, the sum of price elasticities for a given input or output is zero, so that price elasticities for the *numeraire* (K) can easily be recovered. For example, the own-price elasticity of input K would be $\varepsilon_{KK}^{Wr} = -\sum_{k=1}^{K-1} \varepsilon_{Kk}^{Wr}$, where

$$\text{(7)} \qquad \varepsilon_{Kk}^{Wr} = \frac{W_k \times (r_k/r_K)}{W_K} \times \sum_{k'=1}^{K-1} \varepsilon_{kk'}^{Wr}$$

$$k = 1, 2, \ldots, K-1.$$

Crop yield elasticities are also easily computed as the difference between output and land elasticities for the same crop and for a particular price (or subsidy rate).

With the normalized form of profit, the condition of linear homogeneity in prices is automatically satisfied (see Moro and Sckokai 1999). From equation (5), restrictions for land adding-up reduce to the following parametric constraints:

$$\text{(8)} \qquad \sum_{c'=1}^{C} \beta_{cc'}^{p\tau} = \sum_{c'=1}^{C} \beta_{cc'}^{\tau\tau} = \sum_{c'=1}^{C} \beta_{c'k}^{\tau r} = 0$$

$$\forall c, \ \forall k; \quad \sum_{c'=1}^{C} \beta_{c'}^{\tau} = 0;$$

$$\sum_{c'=1}^{C} \gamma_{c'}^{\tau} = 1.$$

To draw valid conclusions regarding production technology from the various elasticities obtained from the optimal farmer's decisions, we need to obtain consistent and efficient parameter estimates. The availability of pooled cross-sections in the agricultural context has led many researchers to perform econometric estimation of the system of equations (4)–(6) using panel data. We will see in the next section that several issues must be dealt with, when considering estimation on a panel of farmers engaged in multi-output activities.

Econometric Considerations

Two major issues require particular attention to estimate the system of equations (4)–(6) using panel data. The first is selection: because of crop rotations and/or short-term changes in crop profitability, a farmer may not plant some crops in a particular year, resulting in a zero value for the corresponding land and output observation. The conditions under which selection occurs must be examined to detect in particular if selection is purely random across farmers, years, and crops. In the case of correlation between the underlying selection process and the production model equations, we need to correct for the resulting selection bias to obtain consistent parameter estimates (Vella 1998). The fact that, with panel data, dynamic patterns in land and production decisions are observed allows for more flexibility in, but also implies a more careful examination of the specification of the selection process.

Second, the distinguishing feature of panel data analysis is to draw inference on parameter estimates based on assumptions about unobserved individual effects. Farmer-specific effects may capture unobserved components related to climate, soils, or farmer's expertise, etc., in the structural equations for output, land, and input. Such effects are also likely to be present in the probability of crop selection and may originate from the same kind of unobserved components. Overlooking individual effects may result in biased parameter estimates if they are correlated with explanatory variables (see Wooldridge 1995).

To deal with these issues, we first represent the optimal production decisions in equations (4)–(6) as a system of equations in which each dependent variable (output, land) can be censored at zero. Letting y_{ict} denote the dependent variable for farmer i, crop c and time t (year), we assume:

$$\text{(9)} \qquad y_{ict} = d_{ict} \times y_{ict}^*, \quad d_{ict} = I(d_{ict}^* > 0)$$

$$\text{(10)} \qquad y_{ict}^* = \boldsymbol{X_{ict}\beta_c} + \alpha_{ic} + \varepsilon_{ict},$$

$$\text{(11)} \qquad d_{ict}^* = \boldsymbol{Z_{ict}\delta_c} + \eta_{ic} + u_{ict}, \ \ i = 1, 2, \ldots, N;$$

$$c = 1, 2, \ldots, C; \ \ t = 1, 2, \ldots, T_i$$

where y_{ict}^* is the latent variable corresponding to output or land, d_{ict}^* is the latent variable for the selection equation, $I(.)$ is the indicator function, d_{ict} is a dummy variable equal to one when $d_{ict}^* > 0$, $\boldsymbol{X_{ict}}$ is the vector of explanatory variables (prices, unit subsidy rates, total land), and $\boldsymbol{Z_{ict}}$ is a vector of explanatory variables with possibly common elements with $\boldsymbol{X_{ict}}$. Since crop selection implies a positive value of output and land, we have $d_{ict} = I(d_{ict}^* > 0) = I(y_{ict}^* > 0)$. α_{ic} and η_{ic} are farmer- and crop-specific individual effects, and ε_{ict} and u_{ict} are independent and identically distributed error terms. We assume that α_{ic} is uncorrelated with ε_{ict}, and that η_{ic} is uncorrelated with u_{ict}.

Selection with Crop Rotations and Unobserved Heterogeneity

As mentioned above, there are several reasons why a crop would not be planted by a farmer in a particular year (i.e., $y_{ict} = d_{ict} = 0$). First, the crop may not be part of the crop rotation for agronomical and/or pest control reasons. When data are not available at the plot level, we may use past and present crop planting decisions to represent crop rotations in the selection equation. Second, a change in relative output prices (or in relative input prices if this crop is requiring relatively more of some inputs than other crops) may make this crop temporarily less attractive. To account for this in the selection equation, it is reasonable to use the set of past prices in equation (11).

As the latent variable d^*_{ict} determines the selection of crop c at date t, past land use decisions are obvious candidates for inclusion in $\boldsymbol{Z_{ict}}$. Let $(s_{i1,t-1}, \ldots, s_{iC,t-1})'$ denote the vector of lagged proportions of land per crop, with $s_{ic,t-1} = \frac{l_{ic,t-1}}{L_{i,t-1}}$. Because $\sum_{c=1}^{C} s_{ict} = 1$, $\forall i, t$, we drop one component, say $s_{iC,t-1}$, and define $\boldsymbol{S_{ict}} = (s_{i1,t-1}, \ldots, s_{ic,t-1}, \ldots, s_{iC-1,t-1})'$. The latent variable d^*_{ict} is then rewritten as

$$(12) \qquad d^*_{ict} = \boldsymbol{S_{ict}\pi_c} + \boldsymbol{\tilde{Z}_{ict}\gamma_c} + \eta_{ic} + u_{ict}$$
$$\boldsymbol{Z_{ict}} = (\boldsymbol{S_{ict}}, \boldsymbol{\tilde{Z}_{ict}}), \boldsymbol{\delta_c} = (\boldsymbol{\pi_c}, \boldsymbol{\gamma_c})$$

where the vector of parameters $\boldsymbol{\pi_c}$ captures the influence of the cropping system on the latent variable d^*_{ict}. The latter can be interpreted as a profitability differential: crop c will be considered at time t only if $d^*_{ict} > 0$. Contemporaneous and farmer-specific prices are not valid candidates for $\boldsymbol{\tilde{Z}_{ict}}$, because corresponding observations would be missing if crop c is not planted in year t. On the other hand, average price and subsidy rates (e.g., at the district level) are admissible, provided all crops are planted in all districts and every year.

Next, consider the issue of unobserved individual heterogeneity which may be correlated with explanatory variables in the selection equation. The farmer- and crop-specific effect η_{ic} can be interpreted as the permanent component picking up the influence of the "average" crop system on the probability of growing this particular crop. To control for a possible correlation between $(\boldsymbol{S_{ict}}, \boldsymbol{\tilde{Z}_{ict}})$ and η_{ic} in (12), we can either project this effect on the whole vector of lags and leads of all explanatory variables $\boldsymbol{Z_{ict}}$ (Wooldridge 1995), or on their individual means (see Papke and Wooldridge 2008 for an application to nonlinear models). The second approach is based on a special case of Chamberlain (1984), such that the regression function of η_{ic} on $\left(\boldsymbol{\bar{S}_{ic}} = \frac{1}{T_i}\sum_{t=1}^{T_i} \boldsymbol{S_{ict}}, \boldsymbol{\bar{\tilde{Z}}_{ic}} = \frac{1}{T_i}\sum_{t=1}^{T_i} \boldsymbol{\tilde{Z}_{ict}}\right)$ is linear:

$$(13) \qquad \eta_{ic} = \boldsymbol{\bar{S}_{ic}\lambda_{1c}} + \boldsymbol{\bar{\tilde{Z}}_{ic}\lambda_{2c}} + \nu_{ic} = \boldsymbol{\bar{Z}_{ic}\lambda_c} + \nu_{ic}$$

where ν_{ic} is a random component and is assumed distributed $N(0, \sigma^2_{\nu,ct})$. We further assume the modified random component in the selection equation, $\kappa_{ict} = \nu_{ic} + u_{ict}$, is independent of $\boldsymbol{\bar{Z}_{ic}}$ and is distributed according to $N(0, \sigma^2_{\kappa,ct})$. Hence, heteroskedasticy (time-dependent and/or farmer-specific) is allowed.

Note that a possible source of endogeneity lies with $\boldsymbol{S_{ict}}$, the only farmer-specific component in the selection equation, which may still be correlated with κ_{ict}. In case endogeneity is not eliminated by the treatment of unobserved individual heterogeneity in equation (13), an instrumental-variable method is available, along the lines of Papke and Wooldridge (2008). In our case, the vector of lagged proportions of land per crop can be instrumented by lagged prices and subsidy rates. A simple exogeneity test can be performed, based on the residuals of the instrumenting equations, which are used as additional explanatory variables in the selection equation.

Dealing with Multivariate Selection

Estimated selection equations can be used for correcting the bias due to multivariate selection of crops. When it is assumed that error terms are only correlated between each pair of selection equation and the corresponding output or land equation, the type 2 Tobit procedure can be used (see Yen 2005). This approach specifies a joint parametric distributional assumption on ε_{ict} and κ_{ict} to evaluate the conditional expectation of the dependent variable given the selection equation. Assuming that ε_{ict} and κ_{ict} are jointly normally distributed and that $E(\alpha_{ic}\boldsymbol{X_{ict}}) = E(\alpha_{ic}\varepsilon_{ict}) = 0$, this conditional expectation is

$$(14)\quad E(y^*_{ict}|\boldsymbol{X_{ict}}, \alpha_{ic}, d^*_{ict} > 0) = \boldsymbol{X_{ict}\beta_c} + \alpha_{ic} + \frac{\sigma_{\varepsilon\kappa,ct}}{\sigma_{\kappa,ct}} \times \frac{\phi[(\boldsymbol{Z_{ict}\delta_c} + \boldsymbol{\bar{Z}_{ic}\lambda_c})/\sigma_{\kappa,ct}]}{\Phi[(\boldsymbol{Z_{ict}\delta_c} + \boldsymbol{\bar{Z}_{ic}\lambda_c})/\sigma_{\kappa,ct}]}$$

where $\sigma_{\varepsilon\kappa,ct}$ and $\sigma_{\kappa,ct}$ respectively denote the covariance between $\boldsymbol{\varepsilon}$ and $\boldsymbol{\kappa}$ and the standard deviation of $\boldsymbol{\kappa}$, and $\phi(\cdot)$ and $\Phi(\cdot)$ are the standard normal density and probability distribution functions. Note that failure of the assumption that $E(\alpha_{ic}\boldsymbol{X_{ict}}) = 0$ can be dealt with by applying a fixed-effect procedure to the equation for y^*_{ict}, as explained in the next section. The probability that $d^*_{ict} > 0$ can also be written

$$(15)\quad \Pr ob(d^*_{ict} > 0) = \Phi\left(\frac{\boldsymbol{Z_{ict}\delta_c} + \boldsymbol{\bar{Z}_{ic}\lambda_c}}{\sigma_{\kappa,ct}}\right) = \Phi_{ct}(\boldsymbol{Z_{ict}\delta_c} + \boldsymbol{\bar{Z}_{ic}\lambda_c})$$

where heteroskedasticity is explicitly accounted for by indexing the probability by time. Once the selection probability (15) is computed (e.g., with the Heckman two-step procedure), estimation is performed on the whole sample by considering

$$(16)\quad E(y^*_{ict}|\boldsymbol{X_{ict}}, \alpha_{ic}) = \Phi_{ct}(\boldsymbol{Z_{ict}\hat{\delta}_c} + \boldsymbol{\bar{Z}_{ic}\hat{\lambda}_c}) \times \left\{(\boldsymbol{X_{ict}\beta_c} + \alpha_{ic}) + \frac{\sigma_{\varepsilon\kappa,ct}}{\sigma_{\kappa,ct}}\phi_{ct} \times (\boldsymbol{Z_{ict}\hat{\delta}_c} + \boldsymbol{\bar{Z}_{ic}\hat{\lambda}_c})\right\} + 0 \times [1 - \Phi_{ct}(\boldsymbol{Z_{ict}\hat{\delta}_c} + \boldsymbol{\bar{Z}_{ic}\hat{\lambda}_c})]$$

where $\boldsymbol{\hat{\delta}_c}$ and $\hat{\lambda}_c$ refer to first-step (Probit) estimates. This type 2 Tobit procedure is based on a single-selection rule (Shonkwiler and Yen 1999; Yen 2005) because only the selection mechanism of the corresponding crop is controlled for. However, restricting the selection rule to the same crop overlooks the fact that selection equations may be correlated through a common unobserved component, and also that the expected output level (or land) for a crop may be correlated with the probability of other crops being planted the same year. This means that the estimation procedure based on equation (16) does not account for the probability that $d^*_{ic't} > 0$ for $c \neq c'$ in the determination of the conditional expectation of y^*_{ict}. As a consequence, sample selection bias is likely if correlation exists between ε_{ict} and $\kappa_{ic't}, c \neq c'$, and if omitted selection terms are correlated with explanatory variables. The source of such correlation may well be crop rotations, which involve specific relationships between land uses across time periods. In any case, the above procedure is inefficient because it does not exploit full information about error correlation (Yen 2005). In such a situation, nonzero correlations between ε_{ict} and $\kappa_{ic't}, c, c' = 1, \ldots, C$, should be considered. In econometric terms, this means that the expectation of a particular dependent variable conditional on all structural equations being positive,

$$(17)\quad E(y^*_{ict}|d^*_{i1t} > 0, \ldots, d^*_{ict} > 0, \ldots, d^*_{iCt} > 0) = E(y^*_{ict}|\kappa_{i1t} > -\boldsymbol{Z_{i1t}\delta_1} - \boldsymbol{\bar{Z}_{i1}\lambda_1}, \ldots \kappa_{ict} > -\boldsymbol{Z_{ict}}\delta_c - \boldsymbol{\bar{Z}_{ic}\lambda_c}, \ldots, \kappa_{iCt} > -\boldsymbol{Z_{iCt}\delta_C} - \boldsymbol{\bar{Z}_{iC}\lambda_C}) \quad \forall c, c' = 1, 2, \ldots, C$$

is in general different from $E(y^*_{ict}|d^*_{ict} > 0)$, even if $E(\kappa_{ict}\kappa_{ic't}) = 0$, $\forall c' \neq c$, unless $E(\varepsilon_{ict}\kappa_{ic't}) = 0$, $\forall c' \neq c$.

A Simple Estimator for Multivariate Selection Rules with Panel Data

To account for the multivariate selection issue, we propose a simple estimation method which generalizes the single-selection rule approach presented above, by extending the panel data approach of Wooldridge (1995) to the case of multivariate selection rules. This extension relies on the assumption that we can find explanatory variables ($\boldsymbol{Z}$) in the selection equations such that the error terms in those equations ($\boldsymbol{\kappa}$) are uncorrelated. On the other hand, correlation pattern are unrestricted between a particular structural equation and the set of selection equations for the same time period.

Formally, we make the following linearity assumption to extend the framework of Wooldridge (1995) to the multivariate selection case:

$$(18)\quad E(\varepsilon_{ict}|\kappa_{ic't}) = \rho_{cc',t} \times \kappa_{ic't} \quad \forall c,\ c' = 1, \ldots, C$$

with ε_{ict} and κ_{ict} jointly normally distributed, and $\rho_{cc',t}$ is a parameter. We add the important restrictions that

$$(19)\quad E(\kappa_{ict}\kappa_{ic't}|\boldsymbol{Z_{ict}},\boldsymbol{Z_{ic't}})=0,\quad \forall c'\neq c$$

$$(20)\quad E(\varepsilon_{ict}\kappa_{ics}|\boldsymbol{Z_{ict}},\boldsymbol{Z_{ics}})=0,\quad \forall s\neq t.$$

Assumption (19) states that there exist two sets of conditioning variables in any pair of selection equations, to ensure that the remaining unobserved heterogeneity terms will be uncorrelated. This implies, in particular, that farmer-specific components such as ν_{ic} and $\nu_{ic'}$ from equation (13) are also uncorrelated for the same individual. In economic terms, crop profitability (of any sign) can be reflected through the latent variables d^*_{ict}; therefore, assumption (19) implies that profitability at any given time period is uncorrelated across crops, once economic factors such as prices and subsidies, and agronomic factors such as lagged proportions of land per crop, are accounted for. Hence, such (observed) factors are the only drivers underlying the decision to grow crops simultaneously. As we will see, assumption (19) is testable.

Assumption (20) is more restrictive than in Wooldridge (1995), as it requires that for a given pair of output or land, and selection equations (for the same c), error terms $\boldsymbol{\kappa}$ and $\boldsymbol{\varepsilon}$ only exhibit contemporaneous correlation. To interpret assumption (20), suppose a random shock (pest, climate) affects both profitability and output negatively at the beginning of a given time period (correlation between $\boldsymbol{\varepsilon}$ and $\boldsymbol{\kappa}$ is positive in this case). Assumption (20) then implies that the farmer will be able to allocate land and inputs such that the random shock on profitability will not affect future output levels. Because the farmer is price-taker, his/her past decisions will be entirely transmitted to future cropping seasons through lagged proportions of land to crop, capturing agronomic conditions. Considering future time periods, we can also interpret this assumption by saying that expectations on future profitability of a crop do not depend on present output or land level which would not be already captured in present prices and subsidies, or in lagged land-use decisions.

We then use the multivariate selection rule of Catsiapis and Robinson (1982), where the dependent variable is observed only if all latent variables underlying the selection equations are positive (such a procedure was also considered by Maddala 1983). Given the assumption of joint normality of the error terms, we have

$$\begin{aligned}(21)\quad & E(y^*_{ict}|\boldsymbol{X_{ict}},\alpha_{ic},\kappa_{i1t}>-\boldsymbol{Z_{i1t}\delta_1}\\ &\quad-\boldsymbol{\bar{Z}_{i1}\lambda_1},\ldots,\kappa_{ict}>-\boldsymbol{Z_{ict}\delta_c}\\ &\quad-\boldsymbol{\bar{Z}_{ic}\lambda_c},\ldots,\kappa_{iCt}>-\boldsymbol{Z_{iCt}\delta_C}\\ &\quad-\boldsymbol{\bar{Z}_{iC}\lambda_C})\\ &=\boldsymbol{X_{ict}\beta_c}+\alpha_{ic}+\sum_{j=1}^{C}\pi_{cj,t}\lambda_{jt}\end{aligned}$$

where $\pi_{cj,t}$ is proportional to $\rho_{cj,t}$ and

$$(22)\quad \lambda_{jt}(\boldsymbol{Z_{ijt}\delta_j}+\boldsymbol{\bar{Z}_{ij}\lambda_j})=\frac{\phi_{jt}(\boldsymbol{Z_{ijt}\delta_j}+\boldsymbol{\bar{Z}_{ij}\lambda_j})}{\Phi_{jt}(\boldsymbol{Z_{ijt}\delta_j}+\boldsymbol{\bar{Z}_{ij}\lambda_j})}.$$

Given the assumption that selection probabilities are uncorrelated, we also have that

$$\begin{aligned}(23)\quad & E(y^*_{ict}|\boldsymbol{X_{ict}},\alpha_{ic})\\ &=\Pr(y^*_{ict}>0|\boldsymbol{Z_{ict}},\eta_{ic})\\ &\quad\times E(y^*_{ict}|y^*_{ict}>0,\boldsymbol{X_{ict}},\alpha_{ic})\\ &\quad+\Pr(y^*_{ict}\leq 0|\boldsymbol{Z_{ict}},\eta_{ic})\times 0\\ &=\Phi_{ct}(\boldsymbol{Z_{ict}\delta_c}+\boldsymbol{\bar{Z}_{ic}\lambda_c})\\ &\quad\times\left(\boldsymbol{X_{ict}\beta}_c+\boldsymbol{\alpha_{ic}}+\sum_{j=1}^{C}\pi_{cj,t}\lambda_{jt}\right).\end{aligned}$$

Note that in equations (22) and (23), we maintain the notation with density and probability distribution functions depending on time and crop, indicating the sample selection correction term is obtained as in Wooldridge (1995), see equation (16).

Estimation of the final model can proceed by estimating equation (21) on the subsample of observations for which $d_{ict}=1$, with a fixed-effects (Within) procedure to eliminate α_{ic}. However, when the panel is severely unbalanced (e.g., when the number of missing observations for some crops is large), we may not have enough noncensored observations remaining from the sample. We can then estimate equation (23) on the whole sample instead, by controlling for fixed effects α_{ic} by a Mundlak or Wooldridge-like procedure. More precisely, we specify

$$(24)\quad \alpha_{ic}=\boldsymbol{\bar{X}_{ic}\mu}_c+\omega_{ic}$$

where we assume ω_{ic} is not correlated with $\bar{\boldsymbol{X}}_{ic} = \frac{1}{T_i}\sum_{t=1}^{T_i}\boldsymbol{X}_{ict}$. In such a case, the model is nonlinear in fixed effects and requires the selection probability to be defined over all observations (to be able to compute selection-correction terms λ_{jt}).

To summarize, our estimation method proceeds in two steps. In the first stage, we estimate the selection equations by Probit, and check that the no-correlation conditions in equation (19) are satisfied. To do this, we run a bivariate Probit on each pair of crops when selection is present, and compute the test statistic for the null hypothesis of no correlation. Determinants of the crop-specific probability include lagged land decisions, current total land, and district-level price indices, and are all defined on the whole sample. As indicated above, exogeneity of lagged proportions of land per crop in the selection equations can be tested for, and if rejected, an instrumental-variable procedure can be applied following Papke and Wooldridge (2008). In the second stage, we estimate the system of structural equations (output, land and input) on the whole sample, according to equation (23). To take care of the possible endogeneity of explanatory variables if some are correlated with individual heterogeneity α_{ic}, we use fixed-effect estimation along the lines of Mundlak, according to equation (24). As in any two-stage estimation method, we need to correct second-stage standard errors of parameter estimates because they depend on first-stage estimates. The computation of a consistent and heteroskedasticity-robust variance-covariance matrix in the case of a first-stage Probit is discussed in Wooldridge (2002, chapter 12). Our procedure can be compared to the single selection approach (type 2 Tobit), which is a special case of our method in which only the index $j = c$ is considered in the sample selection correction term in equation (23). A priori, the single-selection rule will produce biased parameter estimates if omitted correction terms are correlated with explanatory variables. It will be inefficient in any case since the full information about error correlation is not used in estimation.

The Data

The sample consists of 634 farmers selected from the French RICA (FADN, Farm Accounting Data Network) during the period 1995–2001, from three administrative regions composed of twenty-one *départements* (French administrative districts): Pays de Loire (five districts), Midi-Pyrénées (8 districts), and Rhône-Alpes (eight districts). The total number of observations is 2,820. We have selected farmers on the basis of their main production output in arable crops, and do not consider cattle breeding and animal production. Four groups of crops are considered: corn for grain, other cereals (e.g., wheat and rye), oilseeds (e.g., sunflower), and protein crops. We consider five land uses: the same four crops as above, and voluntary land set-aside. The sample is typically unbalanced with respect to years (presence of the farmer in the sample for less than 7 years in total), and crops (land and production are equal to zero for some farmers and years). The average number of years per farmer is 5.27 with a minimum at 3 and a maximum of 7. The proportion of years where corn is selected by a given farmer is 0.66 on average, 0.93 for cereals, 0.72 for oilseed, 0.21 for protein crops, and 0.84 for voluntary land set-aside. The most frequent combinations of crops (in terms of significant estimated correlation coefficients) are cereals-oilseed (0.30), oilseed-protein crops (0.10), and oilseed-set aside (0.35), while the less likely combinations are corn-cereals (−0.11) and corn-oilseed (−0.15).

Output prices for individual crops in each group (corn, cereals, oilseed, and protein crops) are computed by dividing annual crop sales by produced quantity, a unit value approach. The output price index for each group of crops is then computed as a district-level surface-weighted average of unit crop prices. We consider such yearly district averages of output prices for three reasons. First, as discussed above, selection equations should contain nonmissing observations, precluding the use of farmer-specific prices, which would not be observed when the corresponding crop is not planted. Second, using the same definition of prices in all model equations facilitates interpretations. Third, this is a straightforward way to limit the possible measurement-error issue with individual price information.

We construct unit subsidy rates for each crop group and land set-aside from the FADN database, and we compare constructed subsidy rates with DDA (district-level Agricultural Division) data, for the years 1995 to 2001. In practice, some crops may have different subsidy rates depending on irrigation use (Common Agricultural Policy payments are higher in this case). The FADN database does not record irrigation at the crop level for the years considered. We therefore compute an average subsidy rate as the area-weighted average between irrigated and nonirrigated (rainfed) crops. As our main interest lies in production decisions concerning nitrogen fertilizer, we consider only two inputs: chemical fertilizer and other variable inputs. For the latter, we consider an aggregate input consisting of labor, pesticide, seed, irrigation water, and energy. We use the regional, general input price index for agriculture (IPAMPA, Scees) to construct a Stone price index at the *département* (district) level. To do this, yearly average input cost shares (excluding fertilizer) are computed at the district level from FADN, and are used in the Stone price index.

Physical quantities of inputs need to be observed to augment the system of equations (output, land, and input). Because the FADN data do not contain such information for all inputs, we do not consider the aggregate input variable as part of the system of equations, and use the corresponding price as the numeraire. As far as fertilizer input is concerned, we convert fertilizer expenditures to nitrogen fertilizer quantity by using district-level average nitrogen application rates (from cropping practices surveys) and regional-level nitrogen fertilizer unit prices. Consequently, nitrogen fertilizer is the only input considered in our system of equations to be estimated. Average nitrogen fertilizer input per hectare ($2,647/60 = 44.11$ kg N/ha) is particularly large in the three regions considered (see table 1) and, in any case, above the national average nitrogen application rate for total surface on cereals, corn, oilseed, and protein crops.

Finally, crop output $l_c q_c$ and land l_c are obtained directly from the FADN database. Note that, because of the mandatory land set-aside mechanism requiring farmers to "freeze" a minimum proportion of land every year, as a fixed percentage of land allocated to corn, cereals, oilseed, and protein crops, we may conclude a direct relationship between land set-aside and other land areas. This is true for medium- and large-size farms, whereas small

Table 1. Descriptive Statistics for the Sample, 1995–2001

Variable	Mean	Standard deviation
Cereals output (100 kg)	1,546.273	1,816.4451
Corn output (100 kg)	2,213.899	2,866.0331
Protein crop output (100 kg)	384.4246	367.2794
Oilseed output (100 kg)	527.5058	606.7034
Fertilizer input (kg N)	2,647.3060	2,568.3879
Land for cereals (ha)	26.6561	29.7128
Land for corn (ha)	15.8696	26.7715
Land for protein crops (ha)	2.2722	6.0085
Land for oilseed (ha)	15.3844	21.1693
Land for voluntary set-aside (ha)	3.7493	6.0151
Land for cereals (1 if >0)	0.9269	
Land for corn (1 if >0)	0.6702	
Land for oilseed (1 if >0)	0.7223	
Land for protein crops (1 if >0)	0.2088	
Voluntary land set-aside (1 if >0)	0.8386	
Cereals unit price (€/100 kg)	11.8885	2.4999
Corn unit price (€/100 kg)	11.1003	2.2230
Protein crop unit price (€/100 kg)	13.5533	3.1096
Oilseed unit price (€/100 kg)	20.2305	14.8053
Fertilizer unit price (€/kg)	0.5494	0.0585
Cereals unit subsidy (€/ha)	273.4059	28.3150
Corn unit subsidy (€/ha)	357.7379	86.4635
Protein crop subsidy (€/ha)	467.0293	113.6257
Oilseed unit subsidy (€/ha)	479.0407	46.1246
Land set aside subsidy (€/ha)	359.5209	93.4135
Profit (€)	40,177.74	53,987.69

Note: 2,820 observations.

farms are exempted from such requirement. However, a voluntary set-aside system also exists, by which farmers are compensated for hectares set-aside, on top of the mandatory mechanism. To deal with both sources of land set-aside (mandatory and voluntary), we proceed as follows. For observed land set-aside greater than the mandatory rate, we assume that the difference consists in voluntary land set-aside, and retain only this land surface in the land set-aside equation. When observed land set-aside is less than the mandatory rate, we consider that mandatory set-aside scheme does not apply, and the reported land set-aside consists of voluntary set-aside only. Finally, when the proportion of land set-aside is exactly equal to the mandatory rate, we consider that there is no incentive-driven mechanism underlying the farmer's decision, and we set the land set-aside to zero.

The validity of our approach relies on the assumption that conditioning variables $\boldsymbol{Z}_{ict}$ can be found, such that selection equations are not correlated (although they can be correlated with structural equations). Fertilizer input is not censored in our data, so that the multivariate selection rule method only applies to output and land equations. We choose as explanatory variables $\boldsymbol{Z}_{ict}$ the lagged proportions of land to crop and set-aside (excluding the one on c because these proportions sum to 1), district-level output prices, and unit subsidy rates for all crops.

Estimation of the Multi-Output Agricultural Production Model and Elasticities

The first step of our estimation method consists in obtaining crop-specific probability estimates from selection equations, and to test for assumption (19). The probability in equation (15) is year specific to account for heteroskedasticity. To reduce the number of parameters, we estimate instead selection probabilities on two subperiods (1995–1997 and 1998–2001) and perform our specification test on each. As can be seen from table 2, the no-correlation assumption is not rejected for all crop groups and both subperiods. We also test for exogeneity of lagged proportions of land per crop ($\boldsymbol{S}_{ict}$) in the selection equations, by using lagged prices and subsidy rates as instruments. Residuals of the instrumenting equations are incorporated in the land-use Probits, and joint significance of the associated parameters is tested for with a Wald test. Exogeneity is rejected at the 5% level (see table 2) only in the case of oilseed (both subsamples) and protein crops (first subsample); thus, we use fitted values of $\boldsymbol{S}_{ict}$ in these cases.

We then estimate the system of structural equations (output, land, and nitrogen fertilizer equations) by imposing symmetry and land adding-up restrictions, on the whole sample of observations for all land uses. The land adding-up conditions have to be modified slightly however, because total land has to account

Table 2. Test Statistics for Correlation between Selection Equations and Exogeneity of Lagged Land Decisions

Crop	Corn	Cereals	Oilseed	Protein Crop	Land Set-aside
1995–1997					
Corn	–	0.1449 (0.87)	−0.0990 (0.64)	0.0650 (0.78)	0.2824 (0.22)
Cereals		–	−0.2922 (0.55)	0.3691 (0.16)	0.3820 (0.10)
Oilseed			–	−0.1854 (0.40)	−0.1846 (0.72)
Protein crop				–	0.3563 (0.11)
Exogeneity test $\chi^2(4)$	5.20 (0.26)	5.14 (0.27)	9.48 (0.04)	14.30 (0.00)	5.46 (0.24)
1998–2001					
Corn	–	0.2865 (0.09)	−0.1805 (0.33)	−0.1172 (0.37)	−0.0233 (0.91)
Cereals		–	0.0939 (0.72)	0.2398 (0.61)	0.3820 (0.10)
Oilseed			–	−0.1854 (0.40)	0.4482 (0.37)
Protein crop				–	0.2328 (0.52)
Exogeneity test ($\chi^2(4)$)	4.63 (0.32)	5.99 (0.19)	16.10 (0.00)	5.41 (0.24)	4.25 (0.37)

Note: 2,820 observations. Correlation coefficients are obtained from bivariate Probit estimates. For oilseed-protein crops and cereals-land set aside, bivariate Probit estimates were computed on the full sample, as the estimated correlation coefficient were not significantly different across 1995–1997 and 1998–2001 subsamples. The exogeneity test is a Wald test applied to the four residuals from the instrumental equation (one crop dropped due to multicolinearity), distributed as a chi-squared with 4 degrees of freedom under the null assumption of exogeneity. The p-values from the likelihood ratio test of null hypothesis (correlation coefficient equals 0) and the exogeneity test (Wald statistics) are in parentheses.

for mandatory land set-aside, on top of the voluntary one considered in the model. The parametric restrictions becomes

$$(25)\quad \left(\sum_{c=1}^{C'} \beta_c^{\tau}\right)(1+\sigma) + \beta_{set\text{-}aside}^{\tau} = 0$$

$$\left(\sum_{c=1}^{C'} \gamma_c^{\tau}\right)(1+\sigma) + \gamma_{set\text{-}aside}^{\tau} = 1$$

where $C' = $ (corn, cereals, oilseed, protein crops), and σ is the mandatory set-aside rate. As this rate changed every year over the period 1995–2001, the parametric restrictions above are difficult to impose with standard estimation techniques because they would depend on the time period. For this reason, we impose these restrictions using the average land set-aside rate over the years 1995 to 2001.

Imposing convexity of profit with respect to prices is done through a Cholesky decomposition of the Hessian matrix of second-order derivatives of profit, which is semi-definite positive by construction. Diewert and Wales (1987) provide a rank-reduction technique which, by lowering the dimension of the matrix to which the semi-definite positiveness condition is imposed, ensures convexity in prices with no loss of flexibility in the normalized quadratic case. Convergence was achieved by using the rank-reduction technique suggested by Moschini (1998), where the last four columns of the lower diagonal matrix are set to 0.

From our consistent parameter estimates, we compute elasticities from equation (23), accounting not only for the fact that prices (or subsidy rates) also appear in the selection terms, but also in the selection probabilities:

$$(26)\quad \frac{\partial E(y_{ict}|\boldsymbol{X_{ict}})}{\partial p_k} = \Phi_{ict} \times \left(\beta_{ck} + \sum_{c'=1}^{C} \pi_{cc',t} \frac{\partial \lambda_{c't}}{\partial p_k}\right) + \Phi'_{ict}\delta_{ck} \times \left(\boldsymbol{X_{ict}\beta_c} + \boldsymbol{\bar{X}_{ic}\mu_c} + \omega_{ic} + \sum_{c'=1}^{C} \pi_{cc',t}\lambda_{c't}\right)$$

$$= \Phi_{ict} \times \left[\beta_{ck} + \sum_{c'=1}^{C} \pi_{cc',t} \left(\frac{\phi_{ic't}}{\Phi_{ic't}} \times (\boldsymbol{Z_{ic't}\delta_{c'}}) + \left(\frac{\phi_{ic't}}{\Phi_{ic't}}\right)^2\right)\right] + \phi_{ict}\delta_{ck} \times \left(\boldsymbol{X_{ict}\beta_c} + \boldsymbol{\bar{X}_{ic}\mu_c} + \omega_{ic} + \sum_{c'=1}^{C} \pi_{cc',t}\frac{\phi_{ic't}}{\Phi_{ic't}}\right)$$

where $\Phi_{ict} = \Phi_{ct}(\boldsymbol{Z_{ict}\delta_c} + \boldsymbol{\bar{Z}_{ic}\lambda_c})$ and $\phi_{ict} = \Phi'_{ict}$. Note that in the expression above, individual effects α_{ic} have been replaced by the linear projection on the individual means of explanatory variables, according to equation (24). Fixed effects in structural and selection equations are replaced by their estimates according to the Mundlak procedure, and unobserved heterogeneity terms ω_{ic} are set to 0. Elasticities of crop yield are obtained directly as differences between output and land elasticities, and are available from the authors upon request.

Empirical Results

We first present parameter and elasticity of output and land estimates, and we then introduce the environmental simulator used to predict environmental outcomes of production.

Estimation Results

As seen from table 3, all own price elasticities of output and subsidy elasticities of land are significantly different from 0 except for oilseed output, and are below 1 except the own-subsidy elasticity of land for protein crops, which equals 2.30. Furthermore, all significant cross-price elasticities of output and cross-subsidy elasticities of land are negative, except the elasticity of land set-aside with respect to oilseed subsidy. Fertilizer price has a moderate impact on crop output and land, and leads to decrease in yield for all crops except protein crops. This is consistent with agronomic evidence of substitution opportunities between protein crops and nitrogen fertilizer (see Meisinger and Randall 1991). Fertilizer demand has a significant own-price elasticity of −0.37. Other inputs have a low impact on output and land for crops, except land for oilseed (0.12) and land set-aside (1.45), and are complementary to fertilizer (cross-price elasticity of −0.46).

Table 4 reports specification tests for the single-selection rule in output, land, and

Table 3. Elasticities of Output, Land, and Fertilizer: Multivariate Selection Rule Estimation Results

	Price Corn	Price Cereals	Price Oilseed	Price Protein	T Corn	T Cereal	T Oilseed	T Protein	T Set-aside	R Fert	R Others
Output	0.2337[c]	−0.1917[c]	0.0163	0.0027	0.1143[c]	−0.0841[c]	−0.0041	−0.0147	−0.0088	−0.017[c]	0.0402[c]
Corn	(5.15)	(−5.04)	(0.63)	(0.50)	(4.78)	(−5.27)	(−0.14)	(−0.54)	(−1.49)	(−3.93)	(5.24)
Output	−0.1794[c]	0.2380[c]	−0.0185	0.0050	−0.0810[c]	0.0472[b]	−0.0057	0.0409	−0.0053	0.0295[c]	−0.0815[c]
Cereal	(−4.87)	(4.35)	(−0.74)	(0.88)	(−3.84)	(2.32)	(−0.18)	(1.42)	(−0.69)	(3.86)	(−9.35)
Output	0.0395	−0.0369	0.0568	−0.0225[b]	0.0117	0.0579[a]	0.0387	−0.1296[b]	−0.0280[b]	0.0211[a]	0.0073
Oilseed	(0.63)	(−0.59)	(1.59)	(−1.99)	(0.34)	(1.69)	(0.83)	(−2.15)	(−2.06)	(1.71)	(0.87)
Output	0.0763	0.3860	−0.5559[a]	0.3172[c]	−0.0353	−1.0755[c]	−0.1438	1.2825[b]	0.2788[c]	−0.0547	0.0306
Protein	(0.23)	(1.06)	(−1.95)	(2.76)	(−0.15)	(−3.78)	(−0.33)	(2.36)	(2.80)	(−0.52)	(0.28)
Land	0.2064[c]	−0.1730[c]	0.0107	−0.0003	0.1263[b]	−0.0887[c]	−0.0617	0.0483	−0.0174	−0.034[c]	−0.0712
Corn	(4.84)	(−4.43)	(0.42)	(−0.00)	(2.52)	(−2.70)	(−1.42)	(1.04)	(−1.59)	(−2.86)	(−1.16)
Land	−0.1163[c]	0.1197[c]	0.0614[b]	−0.0315[c]	−0.0627[a]	0.1595[c]	0.0629	−0.1873[c]	−0.0333[c]	0.0474[c]	−0.0644[b]
Cereals	(−3.87)	(2.98)	(2.25)	(−3.46)	(−1.79)	(3.51)	(1.47)	(−3.30)	(−2.99)	(3.57)	(−1.95)
Land	−0.0275	−0.0211	0.0201	−0.0035	−0.0667[a]	0.0256	0.1575[a]	−0.1805[b]	0.0215	0.0396[b]	0.1206[a]
Oilseed	(−0.55)	(−0.38)	(0.62)	(−0.29)	(−1.64)	(0.69)	(1.92)	(−2.45)	(1.60)	(2.52)	(1.92)
Land	−0.1698	0.4535	−0.5841[b]	0.2394[b]	0.3118	−1.1749[c]	−1.1947[b]	2.3076[c]	0.0818	−0.391[c]	−3.1515
Protein	(−0.58)	(1.41)	(−2.17)	(2.53)	(1.11)	(−3.76)	(−2.53)	(3.41)	(0.78)	(−4.03)	(−1.12)
Land	−0.1067	−0.03111	−0.1357[b]	0.0565[c]	−0.1518[b]	−0.1513[b]	0.2185[b]	0.1032	0.1234[c]	−0.0163	1.4519[b]
set-as.	(−1.46)	(−0.32)	(−1.96)	(2.80)	(−1.99)	(−2.10)	(2.33)	(0.86)	(3.35)	(−0.50)	(2.36)
X	−0.4073[b]	0.6939[c]	0.1923[a]	−0.0179	−0.4318[c]	0.5573[c]	0.5187[b]	−0.8023[c]	−0.0337	−0.371[b]	−0.0253[c]
Fert.	(−4.17)	(4.02)	(1.71)	(−0.47)	(−2.89)	(3.71)	(2.52)	(−4.05)	(−0.58)	(−1.97)	(−2.79)
X	0.3126[b]	3.1842[c]	−0.1783[c]	−0.0255[b]	−0.0394	0.0140	0.0913[b]	−0.0455	0.0245[b]	−0.195[c]	−0.4644[c]
others	(2.16)	(4.43)	(−3.96)	(−2.36)	(−1.15)	(0.35)	(2.53)	(−1.24)	(2.11)	(−2.98)	(−3.43)

Note: 2,820 observations. Student's *t*-statistics (obtained from heteroskedasticity-robust standard errors corrected for first-stage estimation) are in parentheses. [a], [b] and [c] indicate parameter significant at 10%, 5%, and 1% levels, respectively.

Table 4. Multivariate Selection Rule Estimates: Specification Checks

	q_{corn}	q_{cereal}	$q_{oilseed}$	$q_{protein}$	l_{corn}	l_{cereal}	$l_{oilseed}$	$l_{protein}$	$l_{set-aside}$	$-W_{fert}$
Wald test	117.19	86.37	53.87	67.26	105.11	139.77	79.19	42.72	19.05	120.63
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
R^2	0.5172	0.4192	0.4891	0.5137	0.5040	0.5145	0.5526	0.5348	0.0695	0.1873

Note: 2,820 observations. The Wald test statistic corresponds to the null hypothesis that parameters for Mills ratios are jointly equal to 0 except for the crop considered (p-value in parentheses).

fertilizer equations. Under the null hypothesis of single selection, Mills ratios are jointly equal to 0 except for the crop considered. Wald test statistics for joint parameter significance indicate that the null assumption is rejected in all equations. This is an important empirical result as it confirms that although selection equations are uncorrelated (see table 2), expected output and land (and fertilizer demand) depend significantly on the selection observed for other crops. In other words, the validity of the single rule approach (single Tobit correction for each crop) is strongly rejected in our application, in favour of the multivariate selection rule for correcting multivariate selection bias.

We report in table 5 the own-price and own-subsidy elasticities obtained under alternative specifications from our fixed-effects multivariate selection rule model with lagged proportions of land per crop (Model VI). Model I is estimated on aggregate district-level data without fixed effects and does not need to control for crop selection. Model II specifies random effects instead of fixed effects in the Probit stage, Model III does the same but for the structural output and land equations, Model IV imposes the single selection rule, and Model V omits lagged proportions of land per crop in the selection equations.

Estimating the model on aggregate data (Model I) leads to much higher elasticities in absolute value, in all cases except fertilizer and other inputs, and protein crop output and land. The reduced number of observations does not allow us to consider district-specific fixed effects which may explain why elasticities differ significantly for most output and land variables from the ones obtained with farmer-level data. Indeed, district-specific unobserved components related to climate and soil quality are probably correlated with output and land variables, leading to inconsistent parameter estimates with aggregate data. Moreover, the variation in fertilizer use across districts and years does not result in a significant own-price elasticity.

Compared with Model VI, there is no systematic under- or overestimation of elasticities by Models II to V, but differences in parameter magnitude can be large in either direction. For example, imposing the single selection rule in Model IV leads to underestimation of the protein-crop output elasticity by nearly one third, and the elasticity on land for oilseed by almost one half. Excluding lagged proportions of land per crop from the Probits as in Model V leads in most cases to underestimated parameters in absolute value. The price-elasticity of fertilizer is less affected by alternative specifications, although not considering fixed effects seems to lead to overestimation of the elasticity (in absolute value). Excluding lagged proportions of land per crop (Model V) can lead to very large differences in parameter estimates, for example, 0.01 instead of 0.23 in the case of corn output, or 0.21 instead of 2.30 in the case of land for protein crops. Since our specification is consistent compared to other cases, we cannot conclude unambiguously the direction of the bias. Finally, parameter estimates with our method appear to have similar efficiency (on the basis of standard errors and t-statistics) in seven cases out of eleven.

Application to Environmental Impacts from Nitrogen Fertilizer Use

Nitrate contamination of surface and ground water has become a major environmental issue in many developed countries. Environmental impacts on ground water quality are particularly difficult to evaluate because of lags in the transmission of nitrate from the soil to the water source, and are therefore highly heterogeneous across farmers. Nitrogen runoff is defined as the excess nitrogen available in the soil after harvest, which depends on soil and climate conditions, and on land use, nitrogen input level, and crop yield at the end of the growing season. We consider here a potential pollution indicator for nitrogen emissions in the "below the root" zone. The impact

Table 5. Own-Price Elasticities under Various Model Specifications

	District-Level Aggregate Model I	RE Probit FE Sure MS rule II	FE Probit RE Sure MS rule III	FE Probit FE Sure SS rule IV	FE Probit FE Sure MS rule No lagged land use V	FE Probit FE Sure MS rule VI
Corn output	0.8147[c] (0.2650)	0.0331[c] (0.0118)	0.3386[c] (0.0335)	0.3325[c] (0.0458)[c]	0.0162[c] (0.0127)	0.2337[c] (0.0454)
Cereal output	0.7958[c] (0.1438)	0.2228[c] (0.0631)	0.2545[c] (0.0537)	0.2885[c] (0.0583)	0.1867[c] (0.0509)	0.2380[c] (0.0547)
Oilseed output	0.5477[c] (0.2019)	0.1141[a] (0.0537)	0.0716[c] (0.0282)	0.0662 (0.0440)	0.0959 (0.0589)	0.0568 (0.0358)
Protein crop output	0.2039[a] (0.1239)	0.2206[b] (0.1040)	0.2353[c] (0.0718)	0.1279[a] (0.0773)	0.0549[b] (0.0257)	0.3172[c] (0.1149)
Land for corn	0.6373[b] (0.2816)	0.0391 (0.0367)	0.1330[c] (0.0302)	0.1644[c] (0.0318)	0.0815[b] (0.0308)	0.1263[b] (0.0502)
Land for cereals	0.3775[c] (0.1238)	0.2272[c] (0.0631)	0.1335[c] (0.0349)	0.1339[c] (0.0346)	0.1215[c] (0.0455)	0.1595[c] (0.0455)
Land for oilseed	0.5936[b] (0.2542)	0.3975[b] (0.1512)	0.0688 (0.0621)	0.3058[c] (0.1084)	0.1945[a] (0.1054)	0.1575[a] (0.0821)
Land for protein crops	0.8786[b] (0.3829)	1.5208[c] (0.5222)	1.1319[c] (0.3386)	1.7337[c] (0.5841)	0.2189[b] (0.1096)	2.3076[c] (0.6760)
Land for set-aside	0.2117[a] (0.1279)	0.2595 (0.0505)	0.1507[c] (0.0435)	0.1455[c] (0.0395)	0.1441[c] (0.0391)	0.1234[c] (0.0369)
Fertilizer input	−0.2084 (0.1664)	−0.4455[b] (0.1745)	−0.4728[c] (0.1620)	−0.3620[a] (0.2058)	−0.3266[a] (0.1739)	−0.3712[b] (0.1884)
Other inputs	−0.3083[a] (0.1724)	−0.4512[c] (0.0998)	−0.5254[c] (0.1342)	−0.2070[c] (0.0338)	−0.5728[c] (0.0719)	−0.4644[c] (0.1352)

Note: Robust two-stage corrected standard errors in parentheses except for Model I. FE: Fixed Effects; RE: Random Effects; SS: Single-Selection rule; MS: Multivariate-selection rule. Models II-IV and VI include instrumented lagged land decisions when needed (see text and table 2). [a], [b] and [c] indicate parameter significant at the 10%, 5%, and 1% levels, respectively.

Table 6. Elasticities of Nitrogen Runoff under Various Model Specifications

	District-level aggregate model I	RE probit FE sure MS rule II	FE probit RE sure MS rule III	FE probit FE sure SS rule IV	FE probit FE sure MS rule no lagged land decisions V	FE probit FE sure MS rule VI
Corn price	−0.0195	−0.0865	−0.0087	−0.0438	−0.0718	−0.0845
Cereal price	0.1120	0.1291	0.2450	0.1036	0.2044	0.1773
Oilseed price	0.0476	0.1245	−0.0665	0.0624	0.0311	0.0602
Protein crop price	0.0074	−0.0074	0.0091	−0.0061	−0.0004	−0.0055
Corn subsidy rate	−0.0864	−0.1239	−0.0385	−0.0712	−0.1008	−0.0980
Cereal subsidy rate	0.0923	0.1668	0.1415	0.1067	0.1073	0.1431
Oilseed subsidy rate	0.0124	0.2389	0.0205	0.1796	0.0644	0.1237
Protein crop subsidy rate	−0.0287	−0.2816	−0.1966	−0.2058	−0.0936	−0.2150
Set-aside subsidy rate	−0.0125	−0.0351	−0.0113	−0.0328	−0.0157	−0.0165
Fertilizer price	−0.0854	−0.1407	−0.1420	−0.0974	−0.1120	−0.1204
Other-inputs price	0.0076	−0.1755	−0.1526	−0.1487	−0.1705	−0.1810

Note: Based on 2005 market prices and CAP compensatory-payment unit rates. FE: Fixed Effects; RE: Random Effects; SS: Single-Selection rule; MS: Multivariate-selection rule. Models II-IV and VI include instrumented lagged land decisions when needed (see text and Table 2).

simulator was developed in collaboration with agronomists to conduct a simplified pollution diagnosis on selected geographical areas by mobilising farm-level data (Lacroix et al. 2006). This simulator computes nitrogen runoff from indicators of nitrogen and water balance at the farm level, and is calibrated according to the regional average climate and soil type. Coupling both models is simply achieved by taking land, acreages, crop yields, and fertilizer input at the farm level as inputs to the agronomic simulator. Nitrogen runoff below the root zone (NR) is estimated from experimental results (Lacroix et al. 2006) as:

$$(27) \quad NR = 0.17NB + 33WR + 0.08NB^{*}WR + 0.24IN$$

where WR is the water repletion rate in the soil (ratio between water drained and water kept in the soil), IN is the intercrop duration (the "naked soil" duration), and NB is nitrogen balance (the difference between nitrogen inputs and exports). NB is defined as

$$(28) \quad NB = \sum_c f_c l_c + \sum_j a_j N_j - \sum_c b_c q_c l_c$$

where f_c is the mineral nitrogen application rate per hectare for crop c, l_c is the land allocated to crop c (in hectare), a_j is the average nitrogen supplied by livestock of type j, N_j is the livestock population of type j, b_c is the average nitrogen exported content per unit of yield for crop c, and q_c is the crop yield (per hectare) of crop c. Water balance is combined with soil water capacity and evapo-transpiration, to estimate the winter runoff which ultimately results in nitrogen runoff (equation (27)). Two soil types are considered in the analysis: silt and gravel, with estimated average nitrogen runoff of 55.86 kg N/ha and 107.40 kg N/ha respectively. To save on space, we only report results for silt soil.

We use our estimated elasticities of land choice, crop yield and nitrogen fertilizer to compute the elasticity of the environmental impact (nitrogen runoff) with respect to price and subsidy variables. Although parameters in the production model are estimated over the 1995–2001 period, elasticities are computed locally for prices and CAP compensatory payment unit rates converted to 2005 observed values (see table 6). This corresponds to the last year before partial decoupling was implemented, with the last reform of the CAP. The elasticity of nitrogen runoff is computed over all land uses (corn, cereals, oilseed, protein, crops and land set-aside) and measures the percent change in nitrogen runoff per hectare.

To evaluate the differences between our proposed model and alternative specifications, we also compute the nitrogen runoff elasticities for Models I to V. Consider first elasticities computed with our method (Model VI). All values indicate a fairly inelastic reaction of the nitrogen runoff to prices and subsidy rates (see table 6). This may be due in part to the fact that a change in a particular price or compensatory payment rate is affecting simultaneously land use and crop yield of all crops. It is obviously also due to the time lag between nitrogen fertilizer application and runoff, a minor change in cropping practices (as the one used to compute elasticities) seldomly being sufficient to cause a major change in nitrogen runoff. We therefore expect such an environmental elasticity to be higher in the case of a single-crop model, for which substitution opportunities across crops are not accounted for. The highest average elasticities (in absolute value) are for protein crop subsidy (−0.21), other input price (−0.18), cereal price (0.17), cereal subsidy (0.14), oilseed subsidy (0.12), and fertilizer price (−0.12).

Based on such elasticities, an appropriate way of reducing average nitrogen runoff on the farm would be to reduce cereal and oilseed subsidies, and to increase protein crop subsidy, fertilizer and other input prices. This contradicts common belief that an increase in land for corn (through an increase in corn subsidy) would increase nitrogen runoff and is due to the fact that, while land for corn increases, corn yield decreases. Our analysis reveals that an interesting way to reduce nitrogen runoff is not to target a single crop directly by, e.g., favouring other crops such as cereals as opposed to corn. Furthermore, policies based on taxing nitrogen fertilizer inputs may be partly offset by market conditions for some crops, which are not easily accommodated for by current agricultural policies constantly reducing their degree of intervention on such markets. Moreover, a fertilizer tax policy would imply a very high tax rate to achieve a significant reduction in nitrogen runoff.

Concerning the comparison across model specifications, the model estimated with aggregate data (Model I) yields lower run off elasticities (in absolute value) in all cases except for protein crop price. Hence, while own-price elasticities obtained with Model I were mostly higher than in Model VI, we obtain the opposite effect for nitrogen runoff. Comparing our procedure with single-rule procedures, random-effects or selection equations without dynamics in land shares, no systematic pattern of under- or overestimation seems to emerge. However, differences can be significant across models, particularly with respect to output prices and land subsidy rates. Model V without lagged proportions of land per crop underestimates the absolute value of the runoff elasticity with respect to crop subsidy, compared to the multivariate-selection rule, by a factor of 1.5 to 2. On the other hand, omitting fixed-effects in the selection equations (first-stage Probits) results in higher elasticities (Model II) with respect to subsidies, compared to our procedure. Finally, imposing the single instead of the multivariate selection rule (Model IV) results in underestimation of runoff elasticities for corn and cereal price, subsidy for cereals, and fertilizer price. The conclusion of this exercise is that model specification matters when consistent elasticity estimates are needed for policy assessment in particular. Omitting fixed-effects in various parts of the models, or overlooking multivariate selection of crops, results in our case in significantly different measures of sensitivity of nitrogen runoff to economic variables.

Conclusions

In this article, we propose a method for dealing with multi-output production and land allocation decisions in the presence of corner solutions, i.e., multivariate selection. The estimation procedure we consider deals with unobserved farmers' fixed effects both as part of the crop selection process and in output and land decisions, and it controls for the influence of land-use decisions on the farm at the previous period. We show that a multivariate selection rule should be employed in multi-output settings with corner solutions, instead of a crop-specific type 2 Tobit procedure on the sample of selected crops. Our estimation method requires only panel Probit and simultaneous-equation estimation stages and is able to control for unobserved heterogeneity simultaneously with multivariate selection without using computer-intensive multiple integration techniques. It is consistent as well as efficient, compared to single-selection approaches which do not exploit the full information about error correlation, and which may be subject to selection bias if omitted correction terms are correlated with explanatory variables. The practical limitation, however, is the need to satisfy a zero-correlation condition across selection equations for all crops. This condition is easily testable and was not difficult to satisfy in our application.

The method can be useful in a context of policy evaluation, when the objective is to evaluate the consequence of agricultural policy reforms or the impact of market prices on farmers' output and land decisions. The empirical application on a panel of French farmers reveals that our estimation procedure is preferred to the single selection rule (type 2 Tobit) technique found in the literature on censored systems. We compute elasticities of crop output, land, nitrogen fertilizer, and other inputs and use them as input parameters to an environmental impact simulator for predicting nitrogen runoff.

Comparing estimates obtained from our model with alternative specifications reveals significant differences in run off elasticities. Omitting fixed effects in various parts of the models, or overlooking multivariate selection of crops, results in significantly different measures of sensitivity of nitrogen runoff to economic variables. This result points to the relevance of microeconomic production models estimated with farmer-level data, when production models are considered being used for policy recommendations. Our treatment of crop rotations is not complete however, since farm-level data do not allow us to identify crop rotations over a time horizon sufficient to disentangle the effect of agronomic practices from economic factors (prices, subsidies). A more structural econometric approach of crop rotations with multivariate crop selection is left for future research.

Funding

Financial support from Agence Nationale de la Recherche under project ANR ADD Impacts is gratefully acknowledged.

References

Catsiapis, G., and C. Robinson. 1982. Sample Selection Bias with Multiple Selection Rules. *Journal of Econometrics* 18: 351–368.

Chakir, R., and A. Thomas. 2003. Simulated Maximum Likelihood Estimation of Firms Demand Systems with Corner Solutions and Panel Data: Application to Industrial Energy Demand. *Revue d'Economie Politique* 6: 773–799.

Chamberlain, G. 1984. Panel data. In *Handbook of Econometrics,* ed. Z. Griliches and M. D. Intriligator, 1247–1317. Amsterdam: North-Holland.

Diewert, W. E., and T. J. Wales. 1987. Flexible Functional Forms and Global Curvature Conditions. *Econometrica* 55:43–68.

Fezzi C., and I. J. Bateman. 2009. Structural Agricultural Land Use Modelling. Working Paper, University of East Anglia.

Guyomard, H., M. Baudry, and A. Carpentier. 1996. Estimating Crop Supply Response in the Presence of Farm Programmes: Application to the CAP. *European Journal of Agricultural Economics* 23: 401–420.

Hennessy, D. A. 2006. On Monoculture and the Structure of Crop Rotations. *American Journal of Agricultural Economics* 88: 900–914.

Just, R. E. 1993. Discovering Production and Supply Relationships: Present Status and Future Opportunities. *Review of Marketing and Agricultural Economics* 61: 11–40.

Lacroix, A., F. Laurent, D. Ruelland, and E. Sauboua. 2006. Nitrate Pollution Risk Assessment: From the Model to the Indicator. *International Journal of Agricultural Resources Governance and Ecology* 5: 206–223.

Langpap, C., I. Hascic, and J. Wu. 2008. Protecting Watershed Ecosystems through Targeted Local Land Use Policies. *American Journal of Agricultural Economics* 90: 684–700.

Lee, L.-F., and M.M. Pitt. 1987. Microeconometric Models of Rationing, Imperfect Markets and Non-Negativity Constraints. *Journal of Econometrics* 36: 89–110.

Maddala, G.S. 1983. *Limited–Dependent and Qualitative Variables in Econometrics*. Cambridge, MA: Cambridge University Press.

Meisinger, J. J., and G. W. Randall. 1991. Estimating Nitrogen Budgets for Soil–Crop Systems. In *Managing Nitrogen for Groundwater Quality and Farm Profitability*, ed. R. F. Follett, D. R. Keeney, and R. M. Cruse, 85–124. Madison, WI: Soil Science Society of America.

Millimet, D. L., and R. Tchernis. 2009. Estimating High-Dimensional Demand Systems in the Presence of Many Binding Non-Negativity Constraints. *Journal of Econometrics* 147: 384–395.

Moore, M. R., and D. H. Negri. 1992. A Multi-Crop Production Model of Irrigated Agriculture, Applied to Water Allocation Policy of the Bureau of Reclamation. *Journal of Agricultural and Resource Economics* 17: 29–43.

Moro, D., and P. Sckokai. 1999. Modelling the CAP Arable Crop Regime in Italy: Degree of Decoupling and Impact of Agenda 2000. *Cahiers d'Economie et Sociologie Rurales* 53: 50–73.

Moschini, G. 1998. The Semiflexible Almost Ideal Demand System. *European Economic Review* 42: 349–364.

Mundlak, Y. 1978. On the Pooling of Time Series and Cross Section Data. *Econometrica* 46: 69–85.

Orazem, P. F., and J. A. Miranowski. 1994. A Dynamic Model of Acreage Allocation with General and Crop–Specific Soil Capital. *American Journal of Agricultural Economics* 76: 385–395.

Papke, L. E., and J. M. Wooldridge. 2008. Panel Data Methods for Fractional Response Variables with an Application to Test Pass Rates. *Journal of Econometrics* 145: 121–133.

Perali, F., and J. P. Chavas. 2000. Estimation of Censored Demand Equations from Large Cross–Section Data. *American Journal of Agricultural Economics* 82: 1022–1037.

Plantinga, A. J. 1996. The Effects of Agricultural Policies on Land Use and Environmental Quality. *American Journal of Agricultural Economics* 78: 1082–1091.

Shonkwiler, J. S., and S. T. Yen. 1999. Two–Step Estimation of a Censored System of Equations. *American Journal of Agricultural Economics* 81: 972–982.

Thomas, A. 2003. A Dynamic Model of On–Farm Integrated Nitrogen Management. *European Journal of Agricultural Economics* 30: 439–460.

Vella, F. 1998. Estimating Models with Sample Selection Bias: A Survey. *Journal of Human Resources* 33: 127–169.

Wooldridge, J. M. 1995. Selection Corrections for Panel Data models under Conditional Mean Independence Assumptions. *Journal of Econometrics* 68: 115–132.

Wooldridge, J. M. 2002. *Econometric analysis of cross section and panel data*. Cambridge, MA: MIT Press.

Wu, J., R. M. Adams, C. Kling, and K. Tanaka. 2004. From Microlevel Decisions to Landscape Changes: An Assessment of Agricultural Conservation Policies. *American Journal of Agricultural Economics* 86: 26–41.

Yen, S. T. 2005. A Multivariate Sample–Selection Model: Estimating Cigarette and Alcohol Demands with Zero Observations. *American Journal of Agricultural Economics* 87: 453–466.

Yen, S. T., and B.-H. Lin. 2005. A Sample Selection Approach to Censored Demand Systems. *American Journal of Agricultural Economics* 88: 742–749.

Yen, S. T., B. H. Lin, and D. M. Smallwood. 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: 458–478.