

Panel Data Estimation Techniques for Farm-level Data Model *

Platoni S., Sckokai P. and Moro D.

Istituto di Economica Agro-alimentare, Università Cattolica, 29100 Piacenza, Italy

** This research has been carried out as part of the WEMAC (World Econometric Modelling of Arable Crops) research project (Scientific coordinator: Catherine Benjamin), funded by the European Commission under the 6th Framework programme. The FADN data have been provided in the context of that project.*

Abstract - Econometric models wishing to estimate relevant parameters for agricultural policy analysis are increasingly relying on unbalanced panels of farm-level data. Since in the agricultural economics literature such models have often been estimated through simplified approaches, in this paper we try to verify whether the adoption of more sophisticated panel data techniques may impact the estimation results. For this reason, the policy model by Moro and Sckokai (1999) has been re-estimated using techniques recently developed in the econometric literature. The preliminary results show a strong impact on the estimations. This seems to suggest that the adoption of proper panel-data techniques is likely to be very important in order to obtain reliable estimates of some key policy parameters.

Keywords - Agricultural policy, Panel data, Systems of equations

I. INTRODUCTION

Econometric models wishing to estimate relevant parameters for agricultural policy analysis are increasingly relying on farm-level data, like the Farm Accounting Data Network (FADN - EU) or the Agricultural Resource Management Survey (ARMS - US). The structure of these databases is quite similar, since they are typically unbalanced panels, where we find repeated information on some farms but the same farm may not enter the sample every year. Moreover, they typically collect data referring to a large number of farms, providing very detailed information on farm production activities as well as on farm structural characteristics and resource use.

In recent years, a number of papers have been published drawing relevant policy implications from the estimation of arable crop supply/acreage equations carried out on these databases, either related to the EU Common Agricultural Policy (CAP) (see, e.g., Oude

Lansink and Peerlings (1996), Oude Lansink (1999), Moro and Sckokai (1999), Sckokai and Anton (2005), Serra et al. (2005), Sckokai and Moro (2006) and Serra et al. (2006)) or to the corresponding US policy (see, e.g., Goodwin and Mishra (2006)). However, these papers have always adopted a simplified approach in taking into account the complex econometric issues implied by the use of these databases. In fact, their use implies the adoption of proper panel-data techniques suitable for system of equation estimation, in which the issue of censoring is properly taken into account, since it is very common that not every farm produces each crop every year.

In light of these considerations, the present paper re-examines the analyses proposed for Italy by Moro and Sckokai (1999), adopting a more suitable econometric approach. Thus, we model the CAP arable crop regime using FADN data for Italy in order to analyse supply and acreage response to policy parameters, under the maintained hypothesis of risk-neutral behaviour by farmers. This empirical application has mainly illustrative purposes, since the main objective of the paper is to underline the different results obtained adopting different panel data techniques.

In terms of econometric approach, the paper relies on the Error Component Model (ECM), which is the most frequently used approach to analyse panel data in econometrics. When the panel is incomplete, which is the rule rather than the exception when the data come from large-scale surveys, standard estimation methods cannot be applied (see, e.g., Wansbeek and Kapteyn (1989), Baltagi et al. (2001) and Davis (2002)). Hence the general model we consider is a two-way error component regression for unbalanced panel data, in which both firm and time effects are introduced (among recent empirical applications adopting this approach, see, e.g., Boumahdi et al. (2004)). We

present results obtained using both single equation and system of equation estimation techniques, in which censoring issues have been taken into account using a proper two-step approach.

II. MODEL

A. Theoretical model

The model we adopt refers specifically to the CAP for arable crops as it was implemented before the 2003 reform¹. Under this package, farm income was supported through three main policy tools: the intervention price for cereals, the crop-specific area payments, introduced with the 1993 reform of the CAP, and the compulsory rate of set-aside. Thus, any model wishing to analyse farmers' response to these policy tools have to incorporate them in its assumed decision making structure.

As in Moro and Sckokai (1999), we consider the following profit function for the representative farmer:

$$\begin{aligned} \pi(\mathbf{p}^e, \mathbf{w}, \mathbf{b}, d, s_T, c, \mathbf{z}) \equiv & \max_{\mathbf{y}, \mathbf{x}, s_1, \dots, s_n} \mathbf{p}^e \mathbf{y} - \mathbf{w}' \mathbf{x} + \sum_{i=1}^{n_p} b_i s_i + d s_r \\ \text{s.t. } & \sum_{k=1}^n s_k + s_r \leq s_T \\ & s_r = \frac{c}{1+c} \sum_{i=1}^{n_p} s_i \\ & (\mathbf{y}, \mathbf{x}, \mathbf{z}, \mathbf{s}, s_r) \in T \end{aligned} \quad (1)$$

where \mathbf{y} is the n -dimensional vector of farm outputs and \mathbf{p}^e is the corresponding vector of expected output prices, \mathbf{x} is the m -dimensional vector of variable inputs and \mathbf{w} the corresponding vector of input prices, \mathbf{s} is the vector of land allocations to the n crops, with s_T being total farm land, $n_p < n$ is the number of crops included in the arable crop regime, \mathbf{b} is the vector of crop-specific area payments, d is the set-aside payment, c is the set-aside percentage, s_r is the land that must be set aside, \mathbf{z} is the vector of quasi-fixed inputs in the short run and T is the multi-output short-run technology. Finally, the three constraints are the total land

¹ As it is well known, the most recent reform of the CAP was implemented starting in 2005. Thus, reliable farm-level data referring to the application of the new Single Farm Payment scheme will become available only in the near future.

constraint, the set-aside constraint and the technological constraint, respectively.

If we assume that $\pi(\mathbf{p}^e, \mathbf{w}, \mathbf{b}, d, s_T, c, \mathbf{z})$ is twice continuously differentiable, we can write the following set of derivative properties:

$$\begin{aligned} y_i(\mathbf{p}^e, \mathbf{w}, \mathbf{b}, d, s_T, c, \mathbf{z}) &= \frac{\delta \pi(\cdot)}{\delta p_i^e} \quad i=1, \dots, n \\ x_h(\mathbf{p}^e, \mathbf{w}, \mathbf{b}, d, s_T, c, \mathbf{z}) &= -\frac{\delta \pi(\cdot)}{\delta w_h} \quad h=1, \dots, m \\ s_i(\mathbf{p}^e, \mathbf{w}, \mathbf{b}, d, s_T, c, \mathbf{z}) &= \frac{\delta \pi(\cdot)}{\delta b_i} \quad i=1, \dots, n_p \end{aligned} \quad (2)$$

which allow us to define a set of output supply, input demand and land allocation equations that can be estimated on farm-level data. Since this model estimates simultaneously both supply and land allocation decisions, crop yields become endogenously defined.

B. Empirical specification

For a parametric specification of (1), we rely on the normalized quadratic function, a flexible functional form largely applied to the estimation of agricultural profit functions. This functional form has a Hessian of constants, so the curvature properties can hold globally. Moreover, it allows negative profits, which cannot be managed when logarithmic transformations are used. Choosing p_m^e as the numeraire, the normalized quadratic profit function takes the following general form:

$$\bar{\pi} = a_0 + \mathbf{a}' \bar{\mathbf{q}} + \bar{\mathbf{q}}' \mathbf{A} \bar{\mathbf{q}} \quad (3)$$

where $\bar{\pi} = \frac{\pi}{p_m^e}$, $\bar{\mathbf{q}} = \left(\frac{\mathbf{p}^e}{p_m^e}, \frac{\mathbf{w}}{p_m^e}, \frac{\mathbf{b}}{p_m^e}, s_T, c, \mathbf{z} \right)$ and the scalar a_0 , the vector \mathbf{a} and the matrix \mathbf{A} are parameters to be estimated.

Using the derivative property in (2), output supply, input demand and land allocation equations can be written as:

$$\begin{aligned} y_i &= \left(\alpha_i + \sum_j \alpha_j \bar{q}_j \right) \quad i=1, \dots, n \\ & \quad j=1, \dots, n+m+n_p+2 \\ x_h &= - \left(\beta_h + \sum_j \beta_j \bar{q}_j \right) \quad h=n+1, \dots, n+m-1 \\ & \quad j=1, \dots, n+m+n_p+2 \\ s_i &= \left(\gamma_i + \sum_j \gamma_j \bar{q}_j \right) \quad i=n+m, \dots, n+m+n_p-1 \\ & \quad j=1, \dots, n+m+n_p+2 \end{aligned} \quad (4)$$

where α 's, β 's, and γ 's are appropriate elements of the above vector \mathbf{a} and matrix \mathbf{A} .

Due to the specification of the vector \bar{q} , the homogeneity property is maintained within the empirical model. Moreover, the standard symmetry and reciprocity properties can be imposed with the following parametric restrictions: $\alpha_{ij} = \alpha_{ji}$, $\beta_{ij} = \beta_{ji}$ and $\gamma_{ij} = \gamma_{ji}$.

III. DATA

The data used for the present study are taken from the EU FADN database for the period 1994-2003 (ten years) and refer to the sample of Italian specialised arable crop farms. As mentioned in the introduction, the database is an unbalanced panel of 14,288 individuals observed in the above 10 years period, for a total of 34,140 observations².

The database provides most of the variables needed to estimate the model: crop productions, output prices, land allocations, area payments, family labour, hired labour (number of hours and hourly wages), variable input costs by category (seeds, fertilisers, chemicals, water ...) and quasi-fixed input stock values (buildings and machinery). Variable input prices are not provided by the FADN; thus, price indexes for Italy have been taken from the official Eurostat statistics. The same has been done for deflating capital values, since Eurostat provides also time series of rental price indexes for capital goods.

The initial FADN dataset is very disaggregated, especially in terms of number of outputs and number of variable inputs; thus, to make the estimation feasible, some aggregation has been introduced. We have considered five output categories (durum wheat, maize, oilseeds, other cereals, and other arable crops) with their respective land allocations, where the first

four represent those crops for which the CAP arable crop regime guaranteed different levels of area payments. We have also considered two variable inputs (crop inputs and other variable inputs), one fixed input (total land) and one quasi-fixed inputs (aggregate of capital and family labour). The price of "other inputs" is our numeraire in the normalised quadratic specification. The aggregates have been generated as Laspeyres indexes, while short run profit has been computed as the sum of total gross sales and total area payments minus total variable costs.

Since output prices are unknown at the time land allocation decisions are made, an assumption on how price expectations are formed is needed. We have adopted the well-known "adaptive expectation" hypothesis, following the approach proposed by Chavas and Holt (1990), which implies a correction of lagged prices³. Clearly, since our panel is incomplete, individual (farm) lagged prices cannot be used to construct the series of expected prices. Thus, for each crop, yearly regional average prices have been computed and used to model the mechanism of price expectations⁴.

IV. ECONOMETRIC TECHNIQUES

A. Censoring

As mentioned in the introduction, the estimation of supply and land allocation equations implies the adoption of an appropriate technique to account for censoring, since not every farm produces every crop each year. In order to obtain suitable results for policy analysis, this problem has to be addressed adopting a methodology that uses all the available observations, in order to preserve the representativeness of the FADN sample. For this reason, we used the two-step estimation procedure proposed by Shonkwiler and

² The sample we have used was obtained after the elimination of those farms that presented some "severe outliers" in the key variables needed for the estimation. All farms showing output prices and crop yields falling out of the range defined by the sample mean and two standard deviations were eliminated. The general idea of this procedure is to eliminate those observations that are likely to come from some errors in plugging in the basic data.

³ This correction is based on the assumption that, in each period, farmers update their "naive" expectations (lagged prices) based on the past history of the observed differences between actual prices and "naive" expected prices.

⁴ To avoid the problem of eliminating entire years to model lagged prices, we have used national crop prices taken from Eurostat to model expectations in the first years of our sample.

Yen (1999). Thus, the system of equations in (4) is estimated in the following form:

$$v_{it} = \Phi(\mathbf{h}_{it}, \boldsymbol{\eta}_i^*) f_i(\bar{\mathbf{q}}_{it}, \boldsymbol{\psi}_i) + \rho_i \Theta(\mathbf{h}_{it}, \boldsymbol{\eta}_i^*) \quad (5)$$

where \mathbf{h}_{it} is a vector of variables which explains the binary choice of producing/non producing crop i and $\boldsymbol{\eta}_i^*$ are first-stage probit estimates of the corresponding parameters; v_{it} is any of the dependent variable and $f_i(\cdot)$ is any of the equations of the system in (4); $\boldsymbol{\psi}_i$ is the subset of the normalized quadratic parameters to be estimated that enter equation $f_i(\cdot)$; $\Phi(\cdot)$ and $\Theta(\cdot)$ are the univariate standard normal cumulative distribution and probability density functions, respectively, both computed over probit results, while ρ_i is an extra parameter to be estimated.

The five probit models (one for each output) are estimated using as explanatory variables the level of three quasi-fixed inputs (family labour, buildings and machinery) and one set of dummy variables representing different regions/altitudes⁵. Thus, in each probit model we estimated 10 parameters.

B. Panel data estimation

The panel data estimation relies on the error component model (ECM), which is the most frequently used approach to analyse panel data in econometrics. Since the panel is incomplete, standard estimation methods cannot be applied (see, e.g., Wansbeek and Kapteyn (1989), Baltagi et al. (2001), Davis (2002) and Boumahdi et al. (2004)).

At first, we have estimated our model using the standard single equation one-way fixed effect (FE) and one-way random effect (RE) models, for which estimation commands exist in the most common econometric softwares. As it is well known, one-way FE and RE models assume that differences across individuals can be captured by means of an individual specific intercept term. The FE approach considers

this term as a fixed parameter, while the RE approach considers it as a random disturbance⁶.

In addition, we have estimated a set of two-way RE and FE models, which explore simultaneously both differences across individuals and differences over time for each individual. The econometric software we use (TSP version 5.0) offers the possibility of estimating the Maximum Likelihood (ML) two-way RE model, but we have decided to build our own GLS estimator in order to compute the two-way Hausman test (see Baltagi (2005)).

Note that, when adopting a two-way ECM approach, it is legitimate to consider only the individuals which appear at least twice, since individuals appearing only once do not add any useful information (Wansbeek and Kapteyn (1989)). Thus, after this elimination, our final sample is an unbalanced panel of 7,526 individuals observed in 10 years, for a total of 27,378 observations. In order to compare the results, we have used this reduced database also for one-way FE and RE estimation.

Single equation two-way ECM

The single equation two-way ECM estimation technique for unbalanced panel has been introduced by Wansbeek and Kapteyn (1989) and our estimator has been built following their procedure.

Our unbalanced panel is characterized by a total of n observations, by F farms, indexed by $i=1, \dots, F$, and by T periods, indexed by $t=1, \dots, T$. Let T_i denote the number of times the farm i is observed and F_t the number of farms observed in period t . Hence $\sum_i T_i = \sum_t F_t = n$.

Let \mathbf{D}_i be the $F_t \times F$ matrix obtained from \mathbf{I}_F by omitting the rows corresponding to farms not observed

⁵ Other pertinent variables to be included in probit estimation would be soil quality or demographic characteristics of the farmer (age, education, ...), but unfortunately these variables are not included in the FADN database.

⁶ The one-way FE is the most common approach adopted in agricultural economic studies employing panel data estimation techniques (see, e.g., Oude Lansink and Peerlings (1996), Oude Lansink (1999), Sckokai and Anton, (2005)). Other studies adopt ad hoc simplified approaches, that do not explore specifically the panel structure of the data (see, e.g., Moro and Sckokai (1999), Sckokai and Moro (2006), Serra et al. (2005) and Serra et al. (2006)).

in period t . With $\mathbf{A}_l = (\mathbf{D}'_l \dots \mathbf{D}'_T)'$ and $\mathbf{A}_2 = \text{diag}(\mathbf{D}_l \mathbf{I}_F)$ we can define $\mathbf{A}_F \equiv \mathbf{A}'_l \mathbf{A}_l$, $\mathbf{A}_T \equiv \mathbf{A}'_2 \mathbf{A}_2$ and $\mathbf{A}_{TF} \equiv \mathbf{A}'_2 \mathbf{A}_l$

and we can consider $\mathbf{A} \equiv \begin{pmatrix} \mathbf{A}_l & \mathbf{A}_2 \end{pmatrix}$ which gives the dummy-variable structure for the unbalanced panel model.

In the *two-way FE model* we consider the following matrices:

$$\begin{aligned} \bar{\mathbf{A}} &\equiv \mathbf{A}_2 - \mathbf{A}_l \mathbf{A}_F^{-1} \mathbf{A}'_l = (\mathbf{I}_n - \mathbf{A}_l \mathbf{A}_F^{-1} \mathbf{A}'_l) \cdot \mathbf{A}_2 = (\mathbf{I}_n - \mathbf{P}_{[A_l]}) \cdot \mathbf{A}_2 = \mathbf{Q}_{[A_l]} \mathbf{A}_2 \\ \mathbf{Q} &\equiv \mathbf{A}_T - \mathbf{A}_{TF} \mathbf{A}_F^{-1} \mathbf{A}'_l = \mathbf{A}'_2 (\mathbf{A}_2 - \mathbf{A}_l \mathbf{A}_F^{-1} \mathbf{A}'_l) = \mathbf{A}'_2 \bar{\mathbf{A}} = \mathbf{A}'_2 \mathbf{Q}_{[A_l]} \mathbf{A}_2 \end{aligned} \quad (6)$$

and hence the projection matrix onto the null-space of \mathbf{A} is:

$$\mathbf{Q}_{[A]} = (\mathbf{I}_n - \mathbf{A}_l \mathbf{A}_F^{-1} \mathbf{A}'_l) - \bar{\mathbf{A}} \mathbf{Q}^{-1} \bar{\mathbf{A}}' = \mathbf{Q}_{[A_l]} - \mathbf{Q}_{[A_l]} \mathbf{A}_2 \mathbf{Q}^{-1} \mathbf{A}'_2 \mathbf{Q}_{[A_l]} \quad (7)$$

and therefore the within estimator is:

$$\boldsymbol{\beta}^{WT} = (\mathbf{X}' \mathbf{Q}_{[A]} \mathbf{X})^{-1} (\mathbf{X}' \mathbf{Q}_{[A]} \mathbf{y}) \quad (8)$$

In the *two-way RE model (GLS)* the covariance matrix of the composite error $u_{it} = \mu_i + v_t + \varepsilon_{it}$ is:

$$\boldsymbol{\Omega} = E(\mathbf{u}\mathbf{u}') = \sigma_\varepsilon^2 \mathbf{I}_n + \sigma_\mu^2 \mathbf{A}_l \mathbf{A}'_l + \sigma_v^2 \mathbf{A}_2 \mathbf{A}'_2 \quad (9)$$

With $\tilde{\mathbf{A}}_F = \mathbf{A}_F + \frac{\sigma_\varepsilon^2}{\sigma_\mu^2} \mathbf{I}_F$ and $\tilde{\mathbf{A}}_T = \mathbf{A}_T + \frac{\sigma_\varepsilon^2}{\sigma_v^2} \mathbf{I}_T$, by

defining:

$$\begin{aligned} \tilde{\mathbf{Q}} &= \tilde{\mathbf{A}}_T - \mathbf{A}_{TF} \tilde{\mathbf{A}}_F^{-1} \mathbf{A}'_l \\ \tilde{\mathbf{Q}}_{[A_l]} &= \mathbf{I}_n - \mathbf{A}_l \tilde{\mathbf{A}}_F^{-1} \mathbf{A}'_l \end{aligned} \quad (10)$$

Wansbeek and Kapteyn (1989) show that:

$$\boldsymbol{\Omega}^{-1} = \frac{1}{\sigma_\varepsilon^2} (\tilde{\mathbf{Q}}_{[A_l]} - \tilde{\mathbf{Q}}_{[A_l]} \mathbf{A}_2 \tilde{\mathbf{Q}}^{-1} \mathbf{A}'_2 \tilde{\mathbf{Q}}_{[A_l]}) \quad (11)$$

and then the *GLS* estimator is:

$$\boldsymbol{\beta}^{GLS} = (\mathbf{X}' \boldsymbol{\Omega}^{-1} \mathbf{X})^{-1} (\mathbf{X}' \boldsymbol{\Omega}^{-1} \mathbf{y}) \quad (12)$$

We derive Quadratic Estimations (QUEs) for σ_ε^2 , σ_μ^2 and σ_v^2 by using the FE residuals, averaged over farms or averaged over periods. Since we are considering a constant term, with the FE residuals $\mathbf{e} \equiv \mathbf{y} - \mathbf{X}\boldsymbol{\beta}^{WT}$ and with $\mathbf{f} = \mathbf{E}_n \cdot \mathbf{e} = \mathbf{e} - \bar{\mathbf{e}}$ we equate:

$$\begin{aligned} q_n &= \mathbf{f}' \mathbf{Q}_{[A]} \mathbf{f} \\ q_F &= \mathbf{f}' \mathbf{A}_F \mathbf{A}_F^{-1} \mathbf{A}'_l \mathbf{f} \\ q_T &= \mathbf{f}' \mathbf{A}_T \mathbf{A}_T^{-1} \mathbf{A}'_l \mathbf{f}' \end{aligned} \quad (13)$$

to their expected values:

$$\begin{aligned} E(q_n) &= (n - T - F + I - k) \cdot \sigma_\varepsilon^2 \\ E(q_F) &= (T + k_F - k_0 - I) \cdot \sigma_\varepsilon^2 + \\ &\quad + T \cdot \sigma_\mu^2 + n \cdot \sigma_v^2 - (\lambda_1 \cdot \sigma_\mu^2 + \lambda_2 \cdot \sigma_v^2) \\ E(q_T) &= (F + k_T - k_0 - I) \cdot \sigma_\varepsilon^2 + \\ &\quad + n \cdot \sigma_\mu^2 + F \cdot \sigma_v^2 - (\lambda_1 \cdot \sigma_\mu^2 + \lambda_2 \cdot \sigma_v^2) \end{aligned} \quad (14)$$

with $k_F = \text{tr} \left\{ (\mathbf{X}' \mathbf{Q}_{[A]} \mathbf{X})^{-1} \mathbf{X}' \mathbf{A}_2 \mathbf{A}_T^{-1} \mathbf{A}'_2 \mathbf{X} \right\}$,
 $k_T = \text{tr} \left\{ (\mathbf{X}' \mathbf{Q}_{[A]} \mathbf{X})^{-1} \mathbf{X}' \mathbf{A}_l \mathbf{A}_F^{-1} \mathbf{A}'_l \mathbf{X} \right\}$ and
 $k_0 = \frac{\mathbf{i}'_n \mathbf{X} (\mathbf{X}' \mathbf{Q}_{[A]} \mathbf{X})^{-1} \mathbf{X}' \mathbf{i}_n}{n}$ and with
 $\lambda_1 = \frac{\mathbf{i}'_n \mathbf{A}_l \mathbf{A}'_l \mathbf{i}_n}{n} = \frac{\sum_{i=1}^F T_i^2}{n}$ and $\lambda_2 = \frac{\mathbf{i}'_n \mathbf{A}_2 \mathbf{A}'_2 \mathbf{i}_n}{n} = \frac{\sum_{t=1}^T F_t^2}{n}$.

If $E \left(\begin{matrix} u_{it} \\ x_{it} \end{matrix} \right) \neq 0$ then the estimator $\boldsymbol{\beta}^{GLS}$ becomes biased and inconsistent for $\boldsymbol{\beta}$. Hausman (1978) suggests comparing $\boldsymbol{\beta}^{GLS}$ and $\boldsymbol{\beta}^{WT}$ both of which are consistent under the null hypothesis $H_0: E \left(\begin{matrix} u_{it} \\ x_{it} \end{matrix} \right) = 0$

but which will have different probability limits if H_0 is not true. Baltagi (2005) asserts that the standard Hausman test performs correctly also in the two-way ECM. Therefore we compute:

$$\mathbf{q}_i = \tilde{\boldsymbol{\beta}}^{GLS} - \tilde{\boldsymbol{\beta}}^{WT} \quad (15)$$

where $\tilde{\boldsymbol{\beta}}^{GLS}$ and $\tilde{\boldsymbol{\beta}}^{WT}$ are the vectors of the estimated parameters without the constant term and the Hausman test statistic is given by:

$$m_i = \mathbf{q}_i' \left(\sigma_\varepsilon^2 (\tilde{\mathbf{X}}' \mathbf{Q}_{[A]} \tilde{\mathbf{X}}) - (\tilde{\mathbf{X}}' \boldsymbol{\Omega}^{-1} \tilde{\mathbf{X}}) \right)^{-1} \mathbf{q}_i \quad (16)$$

where $\tilde{\mathbf{X}}$ is the matrix of the regressors without the constant term. Under H_0 the test m_i is asymptotically distributed as $\chi_{\tilde{k}}^2$ where \tilde{k} denotes the dimension of the vector $\tilde{\boldsymbol{\beta}}$.

Two-way SUR system

The most appropriate way of estimating the model in (4) is by a system of equation estimation technique, which in this specific case must be a Seemingly Unrelated Regression (SUR) technique, which allows to impose cross equations restrictions. Once again, standard econometric softwares do not provide automatic commands to estimate two-way SUR systems, since estimating such system implies the adoption of a specific procedure for inverting the variance-covariance matrix of the residuals Ω^{-1} . This procedure has been recently proposed by Biørn (2004) for the estimation of a one-way SUR system. Based on this framework, we have derived the corresponding estimator for the two-way SUR system.

Let consider a system of M equations, indexed by $m=1, \dots, M$. The farms are observed in at least two periods and at most T periods. Let \tilde{F}_p denote the number of farms observed in p periods, with $p=2, \dots, T$. Hence $\sum_p \tilde{F}_p = F$ and $\sum_p \tilde{F}_p p = n$. We assume that the farms are observed in $T-1$ groups such that the \tilde{F}_2 farms observed twice come first, the \tilde{F}_3 farms observed three times come second, etc. Hence with $C_p = \sum_{k=2}^p \tilde{F}_k$ being the cumulated number of farms observed up to p times, the index sets of the farm observed p times can be written as $I_p = \{C_{p-1} + 1, \dots, C_p\}$ where $p=2, \dots, T$ and $C_1 = 0$ (note that I_p may be considered as a pseudo-balanced panel with p observations of each farm).

With k_m being the number of regressors for equation m , the total number of regressors for the system is $k_{SUR} = \sum_{m=1}^M k_m$. Stacking the M equations for the observation (i, t) we have:

$$\mathbf{y}_{it} = \mathbf{X}_{it} \boldsymbol{\beta} + \boldsymbol{\mu}_i + \mathbf{v}_t + \boldsymbol{\varepsilon}_{it} = \mathbf{X}_{it} \boldsymbol{\beta} + \mathbf{u}_{it} \quad (17)$$

where $\mathbf{X}_{it} = \text{diag}(\mathbf{x}_{1it} \dots \mathbf{x}_{Mit})$ and $\boldsymbol{\beta} = \text{diag}(\boldsymbol{\beta}'_1 \dots \boldsymbol{\beta}'_M)'$. With $\boldsymbol{\mu}_i = (\boldsymbol{\mu}_{1i} \dots \boldsymbol{\mu}_{Mi})'$, $\mathbf{v}_t = (\mathbf{v}_{1t} \dots \mathbf{v}_{Mt})'$ and $\boldsymbol{\varepsilon}_{it} = (\boldsymbol{\varepsilon}_{1it} \dots \boldsymbol{\varepsilon}_{Mit})'$, we assume:

$$\begin{aligned} E(\boldsymbol{\mu}_i, \boldsymbol{\mu}_{i'}) &= \sigma_{\mu_{mj}}^2 & i = i' \\ &= 0 & i \neq i' \\ E(\mathbf{v}_t, \mathbf{v}_{t'}) &= \sigma_{v_{mj}}^2 & t = t' \\ &= 0 & t \neq t' \\ E(\boldsymbol{\varepsilon}_{it}, \boldsymbol{\varepsilon}_{i't'}) &= \sigma_{\varepsilon_{mj}}^2 & i = i' \text{ and } t = t' \\ &= 0 & i \neq i' \text{ or } t \neq t' \end{aligned} \quad (18)$$

Hence $\boldsymbol{\mu}_i$, \mathbf{v}_t and $\boldsymbol{\varepsilon}_{it}$ have zero expectations and covariance matrices $\boldsymbol{\Sigma}_\mu$, $\boldsymbol{\Sigma}_v$ and $\boldsymbol{\Sigma}_\varepsilon$. It follows that

$$E\left(\mathbf{u}_{it} \mathbf{u}_{i't'}'\right) = \delta_{i'i'} \boldsymbol{\Sigma}_\mu + \delta_{t't'} \boldsymbol{\Sigma}_v + \delta_{i'i'} \delta_{t't'} \boldsymbol{\Sigma}_\varepsilon \quad \text{with } \delta_{i'i'} = 1 \text{ for } i = i' \text{ and } \delta_{i'i'} = 0 \text{ for } i \neq i', \delta_{t't'} = 1 \text{ for } t = t' \text{ and } \delta_{t't'} = 0 \text{ for } t \neq t'.$$

Let us consider $\mathbf{y}_{i(p)} = \left(\mathbf{y}'_{i1} \dots \mathbf{y}'_{ip} \right)'$,

$\mathbf{X}_{i(p)} = \left(\mathbf{X}'_{i1} \dots \mathbf{X}'_{ip} \right)'$ and $\mathbf{u}_{i(p)} = \left(\mathbf{u}'_{i1} \dots \mathbf{u}'_{ip} \right)'$ for $i \in I_p$ and $p=2, \dots, T$ (and then for $i = 1, \dots, C_2, C_2 + 1, \dots, C_3, \dots, C_{T-1} + 1, \dots, C_T$ with $C_T = F$).

We define the matrix $\mathbf{A}_{i(p)}$ indicating in which period t the farm i is observed. For example with $T=4$ if the farm i is observed in $t=2$ and in $t=4$ we have

$$\mathbf{A}_{i(p)} = \begin{bmatrix} \mathbf{0} & \mathbf{I}_M & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{I}_M \end{bmatrix}. \quad \text{Hence with}$$

$\mathbf{v} = \left(\mathbf{v}'_1 \dots \mathbf{v}'_T \right)'$ for the farm $i \in I_p$ we can define

$\mathbf{v}_{i(p)} = \mathbf{A}_{i(p)} \mathbf{v}$ and we can write the model:

$$\begin{aligned} \mathbf{y}_{i(p)} &= \mathbf{X}_{i(p)} \boldsymbol{\beta} + \left(\mathbf{I}_p \otimes \boldsymbol{\mu}_i \right) + \mathbf{v}_{i(p)} + \boldsymbol{\varepsilon}_{i(p)} = \mathbf{X}_{i(p)} \boldsymbol{\beta} + \mathbf{u}_{i(p)} \\ \boldsymbol{\Omega}_{i(p)} &= \mathbf{I}_p \otimes (\boldsymbol{\Sigma}_\varepsilon + \boldsymbol{\Sigma}_v) + \mathbf{J}_p \otimes \boldsymbol{\Sigma}_\mu = \\ &= \mathbf{E}_p \otimes (\boldsymbol{\Sigma}_\varepsilon + \boldsymbol{\Sigma}_v) + \bar{\mathbf{J}}_p \otimes (\boldsymbol{\Sigma}_\varepsilon + \boldsymbol{\Sigma}_v + p\boldsymbol{\Sigma}_\mu) \end{aligned} \quad (19)$$

Since \mathbf{E}_p and $\bar{\mathbf{J}}_p$ are symmetric and idempotent and have orthogonal columns we have:

$$\boldsymbol{\Omega}_{i(p)}^{-1} = \mathbf{E}_p \otimes (\boldsymbol{\Sigma}_\varepsilon + \boldsymbol{\Sigma}_v)^{-1} + \bar{\mathbf{J}}_p \otimes (\boldsymbol{\Sigma}_\varepsilon + \boldsymbol{\Sigma}_v + p\boldsymbol{\Sigma}_\mu)^{-1} \quad (20)$$

Then we can consider the *GLS* problem for β when Σ_μ , Σ_v and Σ_ε are known, i.e. the problem of minimizing:

$$\sum_{p=2}^T \sum_{i \in I_p} \mathbf{u}'_{i(p)} \mathbf{\Omega}_p^{-1} \mathbf{u}_{i(p)} = \sum_{p=2}^T \sum_{i \in I_p} \mathbf{u}'_{i(p)} \left[\mathbf{E}_p \otimes (\Sigma_\varepsilon + \Sigma_v) \right]^{-1} \mathbf{u}_{i(p)} + \sum_{p=2}^T \sum_{i \in I_p} \mathbf{u}'_{i(p)} \left[\bar{\mathbf{J}}_p \otimes (\Sigma_\varepsilon + \Sigma_v + p\Sigma_\mu) \right]^{-1} \mathbf{u}_{i(p)} \quad (21)$$

If we apply *GLS* on the observations for the farms observed p times we obtain:

$$\begin{aligned} \beta_{kSUR \times 1}^{GLS} &= \left[\sum_{i \in I_p} \mathbf{X}'_{i(p)} \mathbf{\Omega}_p^{-1} \mathbf{X}_{i(p)} \right]^{-1} \times \left[\sum_{i \in I_p} \mathbf{X}'_{i(p)} \mathbf{\Omega}_p^{-1} \mathbf{y}_{i(p)} \right] = \\ &= \left[\sum_{i \in I_p} \mathbf{X}'_{i(p)} \left[\mathbf{E}_p \otimes (\Sigma_\varepsilon + \Sigma_v) \right]^{-1} \mathbf{X}_{i(p)} + \sum_{i \in I_p} \mathbf{X}'_{i(p)} \left[\bar{\mathbf{J}}_p \otimes (\Sigma_\varepsilon + \Sigma_v + p\Sigma_\mu) \right]^{-1} \mathbf{X}_{i(p)} \right]^{-1} \times \\ &\times \left[\sum_{i \in I_p} \mathbf{X}'_{i(p)} \left[\mathbf{E}_p \otimes (\Sigma_\varepsilon + \Sigma_v) \right]^{-1} \mathbf{y}_{i(p)} + \sum_{i \in I_p} \mathbf{X}'_{i(p)} \left[\bar{\mathbf{J}}_p \otimes (\Sigma_\varepsilon + \Sigma_v + p\Sigma_\mu) \right]^{-1} \mathbf{y}_{i(p)} \right] \end{aligned} \quad (22)$$

while the full *GLS* estimator is:

$$\begin{aligned} \beta_{kSUR \times 1}^{GLS} &= \left[\sum_{p=2}^T \sum_{i \in I_p} \mathbf{X}'_{i(p)} \mathbf{\Omega}_p^{-1} \mathbf{X}_{i(p)} \right]^{-1} \times \left[\sum_{p=2}^T \sum_{i \in I_p} \mathbf{X}'_{i(p)} \mathbf{\Omega}_p^{-1} \mathbf{y}_{i(p)} \right] = \\ &= \left[\sum_{p=2}^T \sum_{i \in I_p} \mathbf{X}'_{i(p)} \left[\mathbf{E}_p \otimes (\Sigma_\varepsilon + \Sigma_v) \right]^{-1} \mathbf{X}_{i(p)} + \sum_{p=2}^T \sum_{i \in I_p} \mathbf{X}'_{i(p)} \left[\bar{\mathbf{J}}_p \otimes (\Sigma_\varepsilon + \Sigma_v + p\Sigma_\mu) \right]^{-1} \mathbf{X}_{i(p)} \right]^{-1} \times \\ &\times \left[\sum_{p=2}^T \sum_{i \in I_p} \mathbf{X}'_{i(p)} \left[\mathbf{E}_p \otimes (\Sigma_\varepsilon + \Sigma_v) \right]^{-1} \mathbf{y}_{i(p)} + \sum_{p=2}^T \sum_{i \in I_p} \mathbf{X}'_{i(p)} \left[\bar{\mathbf{J}}_p \otimes (\Sigma_\varepsilon + \Sigma_v + p\Sigma_\mu) \right]^{-1} \mathbf{y}_{i(p)} \right] \end{aligned} \quad (23)$$

We can estimate the covariance matrices Σ_μ , Σ_v and Σ_ε by following either the within-between procedure suggested by Biørn (2004) – corrected for the *two-way model* – or the *QUE* procedure suggested by Wansbeek and Kapteyn (1989) – corrected for the *SUR*.

The first method considers the *FE residuals* $\mathbf{e}_{it} \equiv \mathbf{y}_{it} - \mathbf{X}\beta^{WT}$ for the farm i in period t . If we define $\mathbf{f}_{it} = \mathbf{e}_{it} - \bar{\mathbf{e}}$, the $M \times M$ matrices of within farms, between farms and between times

(co)variations in the \mathbf{f} 's of the different equations are:

$$\begin{aligned} \mathbf{W}_f &= \sum_{i=1}^F \sum_{t=1}^T (\mathbf{f}_{it} - \bar{\mathbf{f}}_i - \bar{\mathbf{f}}_t) (\mathbf{f}_{it} - \bar{\mathbf{f}}_i - \bar{\mathbf{f}}_t)' \\ \mathbf{B}_f^C &= \sum_{i=1}^F \mathbf{T}_i (\bar{\mathbf{f}}_i - \bar{\mathbf{f}}) (\bar{\mathbf{f}}_i - \bar{\mathbf{f}})' \\ \mathbf{B}_f^T &= \sum_{t=1}^T \mathbf{F}_t (\bar{\mathbf{f}}_t - \bar{\mathbf{f}}) (\bar{\mathbf{f}}_t - \bar{\mathbf{f}})' \end{aligned} \quad (24)$$

Since the μ_i 's, the v_t 's and ε_{it} 's are independent we can write:

$$\begin{aligned} \mathbf{W}_f &= E(\mathbf{W}_\varepsilon) \\ \mathbf{B}_f^C &= E(\mathbf{B}_\mu^C) + E(\mathbf{B}_\varepsilon^C) \\ \mathbf{B}_f^T &= E(\mathbf{B}_v^T) + E(\mathbf{B}_\varepsilon^T) \end{aligned} \quad (25)$$

where

$$\begin{aligned} \mathbf{W}_\varepsilon &= \sum_{i=1}^F \sum_{t=1}^T \varepsilon_{it} \varepsilon_{it}' - \sum_{i=1}^F \mathbf{T}_i \bar{\varepsilon}_i \bar{\varepsilon}_i' - \sum_{t=1}^T \mathbf{F}_t \bar{\varepsilon}_t \bar{\varepsilon}_t' \\ \mathbf{B}_\mu^C &= \sum_{i=1}^F \mathbf{T}_i \mu_i \mu_i' - \sum_{i=1}^F \mathbf{T}_i \bar{\mu} \bar{\mu}' = \sum_{i=1}^F \mathbf{T}_i \mu_i \mu_i' - n \bar{\mu} \bar{\mu}' \\ \mathbf{B}_\varepsilon^C &= \sum_{i=1}^F \mathbf{T}_i \bar{\varepsilon}_i \bar{\varepsilon}_i' - \sum_{i=1}^F \mathbf{T}_i \bar{\varepsilon} \bar{\varepsilon}' = \sum_{i=1}^F \mathbf{T}_i \bar{\varepsilon}_i \bar{\varepsilon}_i' - n \bar{\varepsilon} \bar{\varepsilon}' \\ \mathbf{B}_v^T &= \sum_{t=1}^T \mathbf{F}_t v_t v_t' - \sum_{t=1}^T \mathbf{F}_t \bar{v} \bar{v}' = \sum_{t=1}^T \mathbf{F}_t v_t v_t' - n \bar{v} \bar{v}' \\ \mathbf{B}_\varepsilon^T &= \sum_{t=1}^T \mathbf{F}_t \bar{\varepsilon}_t \bar{\varepsilon}_t' - \sum_{t=1}^T \mathbf{F}_t \bar{\varepsilon} \bar{\varepsilon}' = \sum_{t=1}^T \mathbf{F}_t \bar{\varepsilon}_t \bar{\varepsilon}_t' - n \bar{\varepsilon} \bar{\varepsilon}' \end{aligned}$$

Since $E(\mathbf{u}_{it} \mathbf{u}'_{it'}) = \delta_{it'} \Sigma_\mu + \delta_{it} \Sigma_v + \delta_{it'} \delta_{it} \Sigma_\varepsilon$, $E(\mu_i \mu_i') = \delta_{ii} \Sigma_\mu$, $E(v_t v_t') = \delta_{tt} \Sigma_v$ and $E(\varepsilon_{it} \varepsilon'_{it'}) = \delta_{it'} \delta_{it} \Sigma_\varepsilon$ it follows that:

$$\begin{aligned} E(\mathbf{W}_\varepsilon) &= (n - F - T) \cdot \Sigma_\varepsilon \\ E(\mathbf{B}_\mu^C) &= \left(n - \frac{\sum_{i=1}^F \mathbf{T}_i^2}{n} \right) \cdot \Sigma_\mu \\ E(\mathbf{B}_\varepsilon^C) &= (F - 1) \cdot \Sigma_\varepsilon \\ E(\mathbf{B}_v^T) &= \left(n - \frac{\sum_{t=1}^T \mathbf{F}_t^2}{n} \right) \cdot \Sigma_v \\ E(\mathbf{B}_\varepsilon^T) &= (T - 1) \cdot \Sigma_\varepsilon \end{aligned} \quad (26)$$

Hence we can conclude that:

$$\begin{aligned}
\hat{\Sigma}_{M \times M}^{\varepsilon} &= \frac{W_f}{n-F-T} \\
\hat{\Sigma}_{M \times M}^{\mu} &= \frac{\mathbf{B}_f^c - (F-I) \cdot \hat{\Sigma}_{\varepsilon}}{n - \frac{\sum_{i=1}^F T_i^2}{n}} \\
\hat{\Sigma}_{M \times M}^{\nu} &= \frac{\mathbf{B}_f^T - (T-I) \cdot \hat{\Sigma}_{\varepsilon}}{n - \frac{\sum_{t=1}^T F_t^2}{n}}
\end{aligned} \quad (27)$$

The second method considers the *FE residuals* $\mathbf{e}_m \equiv \mathbf{y}_m - \mathbf{X}\boldsymbol{\beta}_m^{WT}$ for the equation $m=1, \dots, M$. If we define $\mathbf{f}_m = \mathbf{E}_n \cdot \mathbf{e}_m = \mathbf{e}_m - \bar{\mathbf{e}}_m$ we can obtain *QUE* for $\sigma_{\varepsilon_{mj}}^2$, $\sigma_{\mu_{mj}}^2$ and $\sigma_{\nu_{mj}}^2$ by equating:

$$\begin{aligned}
q_{n_{mj}} &\equiv \mathbf{f}_m' \mathbf{Q}_{[a]} \mathbf{f}_m \\
q_{F_{mj}} &\equiv \mathbf{f}_m' \mathbf{A}_2 \mathbf{A}_1' \mathbf{A}_2' \mathbf{f}_m \\
q_{T_{mj}} &\equiv \mathbf{f}_m' \mathbf{A}_1 \mathbf{A}_1' \mathbf{A}_1' \mathbf{f}_m
\end{aligned} \quad (28)$$

to their expected values:

$$\begin{aligned}
E(q_{n_{mj}}) &= (n-T-F+I-k_{n_{mj}}-k_j+k_m) \cdot \sigma_{\varepsilon_{mj}}^2 \\
E(q_{F_{mj}}) &= (T+k_{F_{mj}}-k_{0_{mj}}-I) \cdot \sigma_{\varepsilon_{mj}}^2 + (T-\lambda_1) \cdot \sigma_{\mu_{mj}}^2 + (n-\lambda_2) \cdot \sigma_{\nu_{mj}}^2 \\
E(q_{T_{mj}}) &= (F+k_{T_{mj}}-k_{0_{mj}}-I) \cdot \sigma_{\varepsilon_{mj}}^2 + (n-\lambda_1) \cdot \sigma_{\mu_{mj}}^2 + (F-\lambda_2) \cdot \sigma_{\nu_{mj}}^2
\end{aligned} \quad (29)$$

with

$$\begin{aligned}
k_{n_{mj}} &\equiv \text{tr} \left\{ (\mathbf{X}_m' \mathbf{Q}_{[a]} \mathbf{X}_m)^{-1} (\mathbf{X}_m' \mathbf{Q}_{[a]} \mathbf{X}_j) (\mathbf{X}_j' \mathbf{Q}_{[a]} \mathbf{X}_j)^{-1} (\mathbf{X}_j' \mathbf{Q}_{[a]} \mathbf{X}_m) \right\} \\
k_{F_{mj}} &\equiv \text{tr} \left\{ (\mathbf{X}_m' \mathbf{Q}_{[a]} \mathbf{X}_m)^{-1} (\mathbf{X}_m' \mathbf{Q}_{[a]} \mathbf{X}_j) (\mathbf{X}_j' \mathbf{Q}_{[a]} \mathbf{X}_j)^{-1} (\mathbf{X}_j' \mathbf{A}_2 \mathbf{A}_1' \mathbf{A}_2' \mathbf{X}_m) \right\} \\
k_{T_{mj}} &\equiv \text{tr} \left\{ (\mathbf{X}_m' \mathbf{Q}_{[a]} \mathbf{X}_m)^{-1} (\mathbf{X}_m' \mathbf{Q}_{[a]} \mathbf{X}_j) (\mathbf{X}_j' \mathbf{Q}_{[a]} \mathbf{X}_j)^{-1} (\mathbf{X}_j' \mathbf{A}_1 \mathbf{A}_1' \mathbf{A}_1' \mathbf{X}_m) \right\} \\
k_{0_{mj}} &= \frac{\mathbf{1}'_n \mathbf{X}_m (\mathbf{X}_m' \mathbf{Q}_{[a]} \mathbf{X}_m)^{-1} \mathbf{X}_m' \mathbf{Q}_{[a]} \mathbf{X}_j (\mathbf{X}_j' \mathbf{Q}_{[a]} \mathbf{X}_j)^{-1} \mathbf{X}_j' \mathbf{1}_n}{n}
\end{aligned}$$

Since $k_{mj} = k_{jm}$, $k_{F_{mj}} = k_{F_{jm}}$, $k_{T_{mj}} = k_{T_{jm}}$ and $k_{0_{mj}} = k_{0_{jm}}$ obviously we have $\sigma_{\varepsilon_{mj}}^2 = \sigma_{\varepsilon_{jm}}^2$, $\sigma_{\mu_{mj}}^2 = \sigma_{\mu_{jm}}^2$ and $\sigma_{\nu_{mj}}^2 = \sigma_{\nu_{jm}}^2$.

V. RESULTS

In Table 1 we report the estimated parameters for the durum wheat supply equation, one of the most important arable crop in Italy (for space reasons, we cannot report all the estimated parameters). As one can easily appreciate, the adoption of different estimation techniques implies obtaining quite different results,

both in terms of absolute value of the estimated parameters and in terms of their statistical significance⁷.

For example, the own price-response of durum wheat (the p_1 row in Table 1) is significant only in the one-way FE and RE models and in the two-way SUR system, while in other models is not statistically significant. Moreover, among significant parameters, we observe quite a strong variability, since the two-way SUR system estimates a parameter that is approximately 50% higher than those estimated with single equation techniques. The same happens for the other key parameter of the own area payment effect (the b_1 row in the same table). Here all models provide positive and significant parameters, but their absolute value is strongly different among models. Nevertheless the two-way SUR provides the highest value.

In Table 2 we can appreciate how the SUR technique allows the cross equations restrictions $\alpha_{ij} = \alpha_{ji}$ and $\gamma_{ij} = \gamma_{ji}$.

In Table 3 we provide own-price and own-payment elasticities for both the supply and land allocation equations of our model, the key parameters that are used for policy simulations. Once again, results turn out to be quite different across models. For example, limiting our attention to those elasticities significantly different from 0, we have that the own payment elasticity for oilseeds supply (rows y_3 in the table) and the own price elasticity for maize land allocation (row s_2 in table) are much higher in the case of the two-way SUR system as compared to all the other models.

VI. CONCLUDING REMARKS

In recent years, a number of agricultural economics papers have been published drawing relevant policy implications from the estimation of arable crop supply/acreage equations carried out on farm-level data.

⁷ In Tables 1, 2 and 3 we adopt the following convention: y = output supply, p = output price, x = input demand, w = input price, s = land allocation, b = area payment, l = Durum wheat, 2 = Maize, 3 = Oilseeds, 4 = Other Cereals, 5 = Other output, s_T = total land, c = set-aside percentage and z = quasi-fixed input (aggregate of capital and family labour).

However, these papers have often adopted a simplified approach in taking into account the complex econometric issues implied by the use of these data, which are typically unbalanced panels. In fact, the use of these data implies the adoption of proper panel-data techniques, which have been recently developed in the econometric literature and that have still to be incorporated as automatic commands in the standard econometric softwares.

In light of these considerations, the present paper re-examines the analyses proposed for Italy by Moro and Sckokai (1999), adopting a more suitable econometric approach. Thus, we model the CAP arable crop regime using FADN data for Italy in order to analyse supply and acreage response to policy parameters. Our empirical application has mainly illustrative purposes, since the main objective of the paper is to underline the different results obtained adopting different panel data techniques.

In terms of econometric approach, the paper relies on the Error Component Model, both in its one-way version (i.e. considering only the individual specific effect) and in its two-way version (i.e. considering both the individual and the time specific effects). Since the estimated model is a set of simultaneous

equations, the corresponding regressions have been estimated both as single equations and as a SUR system of equations. In adopting this last technique, we have extended the one-way SUR technique proposed by Bjorn (2004) to the two-way case.

The preliminary results of our work confirm our initial expectations, since the adoption of different estimation techniques implies obtaining quite different results, both in terms of absolute value of the estimated parameters and in terms of their statistical significance. This seems to suggest that the adoption of proper panel-data techniques is likely to be very important in order to obtain reliable estimates of some key policy parameters, like the output price and area payment elasticities estimated in our model.

However, in this research a further step is needed, since at this stage we are not able to select, through an appropriate test, among our alternative econometric specifications. In fact, as noted by Wooldridge (2001), in presence of heteroscedasticity the standard Hausman test does not perform correctly. Since heteroscedasticity is very common in farm-level data, and it is present also in our sample, our two-way estimators must be modified in order to account for heteroscedastic error terms.

Table 1 Durum wheat supply estimated parameters under different panel-data techniques

	FE one-way	FE two-way	RE one-way (GLS)	RE one-way (ML)	RE two-way (GLS)	RE two-way (ML)	SURWB two-way (GLS)	SUR QUE two-way (GLS)
<i>cons</i>			113.6480*** (3.4729)	113.1820*** (3.6027)	110.6890*** (4.0557)	112.8970*** (3.6489)	-19.7823*** (1.6353)	-21.3672*** (1.5650)
p_1	0.7387** (0.3162)	-0.3754 (.3816)	0.6600* (0.3048)	0.6598* (0.3078)	-0.2437 (0.3651)	-0.2403 (0.3668)	1.0514*** (0.2914)	1.2239*** (0.2867)
p_2	0.5820*** (0.1807)	0.9533*** (0.1868)	0.4366* (0.1687)	0.4411* (0.1707)	0.7858** (0.1769)	0.7726** (0.1756)	0.1928 (0.1557)	-0.1637 (0.1522)
p_3	-0.6987*** (0.1676)	-0.7061*** (0.1685)	-0.3362** (0.1550)	-0.3569** (0.1570)	-0.4291*** (0.1591)	-0.3529** (0.1573)	-0.3675*** (0.1230)	0.1111 (0.1214)
p_4	-1.3836*** (0.1441)	-1.3270*** (0.1447)	-1.2426*** (0.1300)	-1.2476*** (0.1318)	-1.2118*** (0.1341)	-1.1916*** (0.1319)	-1.1539*** (0.1130)	-1.6945*** (0.1094)
p_5	-0.1466 (0.1329)	-0.3742** (0.1678)	-0.4464*** (0.1265)	-0.4294*** (0.1279)	-0.6008*** (0.1579)	-0.6780*** (0.1584)	-0.6019*** (0.1184)	-0.7146*** (0.1156)
w	-4.4451*** (1.267)	-2.4421* (1.3965)	-7.6837*** (1.2187)	-7.4735*** (1.2322)	-4.9683*** (1.3469)	-5.6935*** (1.3493)	-8.1235*** (1.1586)	-7.5284*** (1.1354)
b_1	0.4052*** (0.0934)	0.4515*** (0.0942)	1.0263*** (0.0856)	0.9928*** (0.0872)	0.9040*** (0.0883)	1.0248*** (0.0875)	1.1202*** (0.0817)	1.1404*** (0.0800)
b_2	-0.0491 (0.0569)	-0.0499 (0.0604)	-0.0537 (0.0536)	-0.0531 (0.0542)	-0.0679 (0.0574)	-0.0747 (0.0571)	-0.1387*** (0.0424)	-0.1365*** (0.0419)
b_3	-0.0536 (0.0342)	0.0845 (0.0402)	-0.0996*** (0.0330)	-0.0975*** (0.0334)	0.0459 (0.0385)	0.0434 (0.0390)	-0.0879*** (0.0293)	-0.1143*** (0.0290)
b_4	0.0434 (0.1952)	-0.0280 (0.2028)	-0.2085 (0.1717)	-0.2052 (0.1742)	-0.2749 (0.1830)	-0.2899 (0.1793)	0.4492*** (0.1015)	0.4068*** (0.1004)
s_r	0.5398*** (0.0125)	0.5369*** (0.0125)	0.4174*** (0.0073)	0.4203*** (0.0076)	0.4312*** (0.0081)	0.4190*** (0.0076)	0.4457*** (0.0069)	0.4309*** (0.0068)
c	-37.0379*** (2.4832)	-37.3892*** (2.4926)	-23.6270*** (2.2858)	-24.4107*** (2.3250)	-27.4723*** (2.3467)	-24.6105*** (2.3273)	-28.4253*** (2.1697)	-27.7734*** (2.1130)
z	0.0011*** (0.0003)	0.0011*** (0.0003)	0.0017*** (0.0002)	0.0016*** (0.0002)	0.0013*** (0.0003)	0.0016*** (0.0002)	0.0021*** (0.0002)	-0.0002 (0.0002)

Standard errors in brackets. *** 1% significance, **5% significance, *10% significance

Table 2 Estimated parameters under two-way SUR QUE (GLS)

	Dependent Variables									
	y_1	y_2	y_3	y_4	y_5	w	s_1	s_2	s_3	s_4
<i>cons</i>	-21.3672*** (1.5650)	-41.8531*** (4.5059)	22.3254*** (5.3982)	-53.5738 (0)	-363.5020*** (38.2655)	-19.1745*** (2.7951)	4.81388*** (0.1986)	-0.1441 (0.1317)	0.3524*** (0.0967)	0.5997*** (0.1563)
p_1	1.2239*** (0.2866)					-1.0193 (0.7681)	0.0110 (0.0543)	-0.0116 (0.0204)	-0.0359 (0.0301)	-0.2858 (0)
p_2	-0.1637 (0.1522)	1.6954*** (.3906)				0.7794* (0.4575)	-0.0451 (0.0317)	0.0854*** (0.0200)	-0.0678*** (0.0207)	0.0644*** (0.0155)
p_3	0.1111 (0.1214)	-0.9397*** (0.1240)	5.3231*** (0.1023)			0.5690 (0.3724)	-0.1364*** (0.0240)	-0.1588*** (0.0109)	1.3750*** (0.0184)	-0.1840 (0)
p_4	-1.6945*** (0.1094)	0.4467*** (0.1549)	-1.5350*** (0.0781)	5.4384*** (0.1317)		1.4904*** (0.3311)	-0.3613*** (0.0223)	-0.0685*** (0.0104)	-0.2250*** (0.0145)	0.8912*** (0.0133)
p_5	-0.7146*** (0.1156)	0.5256** (0.2488)	-0.2004** (0.0936)	-0.0158 (0.1242)	-3.1651*** (1.0924)	0.0277 (0.3308)	-0.0705*** (0.0222)	0.0001 (0.0138)	-0.0178 (0.0158)	0.0428*** (0.0150)
w	-7.5284*** (1.1354)	-3.9609** (1.9462)	2.0045*** (0.7617)	-1.9126* (0.9835)	0.2082 (6.8105)	-5.9019* (3.2356)	-0.0731 (0.2236)	0.0568 (0.1183)	0.2691** (0.1308)	0.0634*** (0.0112)
b_1	1.1404*** (0.0800)	1.3579*** (0.1699)	-0.1391*** (0.0467)	0.1322 (0.0876)	2.8923** (0.7058)	0.0676 (0.2317)	0.1233*** (0.0165)			
b_2	-0.1365*** (0.0419)	0.1276 (0.1411)	-0.0816** (0.0331)	-0.1496*** (0.0515)	-1.4348*** (0.4673)	0.0334 (0.1475)	-0.0018 (0.0074)	0.0231*** (0.0067)		
b_3	-0.1143*** (0.0290)	0.1124 (0.0712)	0.0788*** (0.0272)	0.0353 (0.0330)	0.3742 (0.3027)	-0.1765** (0.0882)	-0.0303*** (0.0054)	-0.0012 (0.0036)	0.0059 (0.0043)	
b_4	0.4068*** (0.1004)	-2.8680*** (0.2481)	0.3155*** (0.0658)	0.0027 (0.1524)	2.7314*** (0.9696)	-1.9420*** (0.3752)	0.0073 (0.0103)	-0.0043 (0.0061)	-0.0008 (0.0040)	-0.0932*** (0.0177)
s_T	0.4309*** (0.0068)	0.3151*** (0.0146)	0.2189*** (0.0034)	0.4903*** (0.0088)	2.8609*** (0.0385)	-1.8111*** (0.0197)	0.0896*** (0.0013)	0.0373*** (0.0009)	0.0550*** (0.0008)	0.0646*** (0.0010)
c	-27.7734*** (2.1130)	37.0367*** (5.1371)	-7.5390*** (1.7587)	-7.0541*** (2.3491)	-59.9017*** (19.4647)	46.2878*** (6.0859)	-8.0540*** (0.4262)	1.3302*** (0.2800)	-1.9134*** (0.2974)	-1.0590*** (0.2889)
z	-0.0002 (0.0002)	0.0181*** (0.0005)	-0.0026*** (0.0001)	-0.0010*** (0.0002)	0.0178*** (0.0017)	-0.0216 (0.0006)	0.0001*** (0.00004)	0.0009*** (0.00003)	-0.0003*** (0.00002)	-0.0001*** (0.00003)

Standard errors in brackets. *** 1% significance, **5% significance, *10% significance

Table 3 Estimated own-price and own-payment elasticities under different panel-data techniques

		FE one-way	FE two-way	RE one-way (GLS)	RE one-way (ML)	RE two-way (GLS)	RE two-way (ML)	SURWB two-way (GLS)	SUR QUE two-way (GLS)
y_1	p_1	0.1206**	-0.0613	0.1077**	0.1077**	-0.0398	-0.0392	0.1716***	0.1998***
	b_1	0.1571***	0.1751***	0.3980***	0.3850***	0.3506***	0.3974***	0.4344***	0.4422***
y_2	p_2	0.0280	0.0650	0.0650	0.0650	0.0766*	0.0848*	-0.0795*	0.1736***
	b_2	0.0599**	-0.0259	0.0886***	0.0886***	0.0053	0.0219	0.0739***	0.0254
y_3	p_3	0.7875***	0.7712***	0.8775***	0.8607***	0.8633***	0.8459***	0.8554***	0.7351***
	b_3	-0.0027	0.0161	0.0561	0.0486	0.0777	0.0349	0.0764	0.1460***
y_4	p_4	0.5797***	0.5806***	0.5774***	0.5765***	0.5753***	0.5767***	0.5837***	0.5738***
	b_4	-0.0624	-0.0773	0.0942*	0.0649	0.0494	0.0524	-0.2021***	0.0009
s_1	p_1	-0.0353	-0.0449	-0.0158	-0.0179	-0.0227	-0.0209	-0.0068	0.0055
	b_1	0.0275	0.0328	0.2242***	0.1979***	0.1956***	0.2012***	0.1663***	0.1467***
s_2	p_2	0.0372	0.0298	0.0584**	0.0574	0.0472*	0.0482**	-0.0079	0.0896***
	b_2	0.0333**	-0.0188	0.0435***	0.0430***	0.0037	0.0016	0.0694***	0.0471***
s_3	p_3	0.4598***	0.4478***	0.5129***	0.5039***	0.5026***	0.4925***	0.5137***	0.4850***
	b_3	0.0483**	0.0242	0.0494***	0.0503**	0.0369	0.0164	0.0334	0.0281
s_4	p_4	0.3719***	0.3713***	0.3886***	0.3831***	0.3786***	0.3819***	0.4080***	0.4187***
	b_4	-0.0623**	-0.0728**	-0.0040	-0.0004	-0.0145	-0.0105	-0.1602***	-0.1371***

*** 1% significance, **5% significance, *10% significance

REFERENCES

1. Baltagi B H (2005) *Econometric analysis of panel data* (3rd edition). Wiley and Sons, Chichester
2. Baltagi B H, Bresson G, Pirotte A (2005) Adaptive estimation of heteroskedastic error component models. *Econometric Reviews* 24: 39-58
3. Baltagi B H, Song S H, Jung B C (2001) The unbalanced nested error component regression model. *Journal of Econometrics* 106: 357-381
4. Biørn E (2004) Regression systems for unbalanced panel data: a stepwise maximum likelihood procedure. *Journal of Econometrics* 122: 281-291
5. Boumahdi R, Chaaban J, Thomas A (2004) Import demand estimation with country and products effects: application of multi-way unbalanced panel data models to Lebanese imports. *Cahier de Recherche INRA de Toulouse* 2004-17
6. Chavas J P, Holt M T (1990) Acreage decision under risk: the case of corn and soybeans. *American Journal of Agricultural Economics* 72: 529-538
7. Davis P (2002) Estimating multi-way error components models with unbalanced data structures. *Journal of Econometrics* 106: 67-95
8. Goodwin B K, Mishra A K (2006) Are 'decoupled' farm program payments really decoupled? An empirical evaluation. *American Journal of Agricultural Economics* 88: 73-89
9. Greene W (2003) *Econometric analysis* (5th edition). Prentice Hall, New York
10. Hausman J A (1978) Specification tests in Econometrics. *Econometrica* 46: 1251-1271
11. Li Q, Stengos T (1994) Adaptive estimation in the panel data error component model with heteroskedasticity of unknown form. *International Economic Review* 35: 981-1000
12. Mazodier P, Trognon (1978) Heteroskedasticity and stratification in error components models. *Annales de l'Insee* 30-31: 451-482
13. 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* 53: 49-73
14. Oude Lansink A, Peerlings J (1996) Modelling the new EU cereals regime in the Netherlands. *European Review of Agricultural Economics* 23: 161-178
15. Oude Lansink A (1999) Area allocation under price uncertainty on Dutch arable farms. *Journal of Agricultural Economics* 50: 93-105
16. Randolph W C (1988) A transformation for heteroscedastic error components regression models. *Economic Letters* 27: 349-354
17. Roy N (2002) Is adaptive estimation for panel model with heteroskedasticity in the individual specific error component? Some Monte Carlo evidence. *Econometric Reviews* 21: 189-203
18. Sckokai P, Anton J (2005) The degree of decoupling of area payments for arable crops in the European Union. *American Journal of Agricultural Economics* 87: 1220-1228
19. Sckokai P, Moro D (2006) Modeling the reforms of the Common Agricultural Policy for arable crops under uncertainty. *American Journal of Agricultural Economics* 88: 43-56
20. Serra T, Zilberman D, Goodwin B K, Featherstone A (2006) Effects of decoupling on the mean and variability of output. *European Review of Agricultural Economics* 33: 269-288
21. Serra T, Zilberman D, Goodwin B K, Hyvonen K (2005) Replacement of agricultural price supports with area payments in the European Union and the effects on pesticide use. *American Journal of Agricultural Economics* 87: 870-884
22. Shonkwiler J S, Yen S T (1999) Two-step estimation of a censored system of equations. *American Journal of Agricultural Economics* 81: 972-982
23. Wansbeek T, Kapteyn A (1989) Estimation of the error-components model with incomplete panels. *Journal of Econometrics* 41: 341-361
24. Wooldridge J M (2001) *Econometric analysis of cross section and panel data*. The MIT Press, Cambridge

Corresponding author:

- Paolo Sckokai
- Istituto di Economia Agro-alimentare
Università Cattolica del Sacro Cuore
via Emilia Parmense 84
29100 Piacenza
Italy
paolo.sckokai@unicatt.it