FARM-LEVEL DATA MODEL FOR AGRICULTURAL POLICY ANALYSIS: A TWO-WAY ECM APPROACH

Paolo Sckokai, Daniele Moro and Silvia Platoni

Istituto di Economia Agro-alimentare
Università Cattolica
Via Emilia Parmense, 84
29100, Piacenza
ITALY
(e-mail: paolo.sckokai@unicatt.it)

Paper prepared for presentation at the 107th EAAE Seminar *"Modelling of Agricultural and Rural Development Policies"*. Sevilla, Spain, January 29th -February 1st, 2008

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.

Key words: *Agricultural policy, Panel data, Systems of equations*

1. Introduction

Econometric models wishing to estimate relevant parameters for agricultural policy analysis are increasingly relying on farm-level data, like the European Union (EU) Farm Accounting Data Network (FADN) or the United States (US) Agricultural Resource Management Survey (ARMS). 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) (Oude Lansink and Peerlings, 1996; Oude Lansink, 1999; Moro and Sckokai, 1999; Sckokai and Anton, 2005; Serra et al., 2005; Sckokai and Moro, 2006; Serra et al., 2006) or to the corresponding US policy (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., Wansbeed 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. Bhoumahdi 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.

2. Model

2.1 Theoretical model

The model we adopt refers specifically to the CAP for arable crops as it was implemented before the 2003 reform[1]. 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:

$$\pi(\mathbf{p^e}, \mathbf{w}, \mathbf{b}, d, s_T, c, \mathbf{z}) \equiv \max_{\mathbf{y}, \mathbf{x}, s_1, \ldots, s_n} \mathbf{p^e}^T \mathbf{y} - \mathbf{w}^T \mathbf{x} + \sum_{i=1}^{n_p} b_i s_i + d s_r$$

(1)

$$s.t. \ \sum_{k=1}^{n} s_k + s_r \leq s_T \qquad s_r = \frac{c}{1-c} \sum_{i=1}^{n_p} s_i \qquad (\mathbf{y}, \mathbf{x}, \mathbf{z}, \mathbf{s}, s_T) \in T$$

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 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:

$$y_i(\mathbf{p^e}, \mathbf{w}, \mathbf{b}, d, s_T, c, \mathbf{z}) = \partial \pi(.)/\partial p_i^e \qquad i = 1, \ldots, n$$

(2) $$x_j(\mathbf{p^e}, \mathbf{w}, \mathbf{b}, d, s_T, c, \mathbf{z}) = -\partial \pi(.)/\partial w_j \qquad j = 1, \ldots, m$$

$$s_i(\mathbf{p^e}, \mathbf{w}, \mathbf{b}, d, s_T, c, \mathbf{z}) = \partial \pi(.)/\partial b_i \qquad i = 1, \ldots, n_p$$

---

[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.

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.

2.2 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_n^e as the numeraire, the normalized quadratic profit function takes the following general form:

(3) $\qquad \overline{\pi} = a_0 + \mathbf{a}^T \overline{\mathbf{q}} + \overline{\mathbf{q}}^T \mathbf{A} \overline{\mathbf{q}}$

where $\overline{\pi} = \pi / p_n^e$, $\overline{\mathbf{q}} = (\mathbf{p^e} / p_n^e, \mathbf{w} / p_n^e, \mathbf{b} / p_n^e, d / p_n^e, s_T, c, \mathbf{z})$ and the scalar a_0, the vector $\mathbf{a}$ and the matrix $\mathbf{A}$ are parameters to be estimated.

Using the derivative property in (5), output supply, input demand and land allocation equations can be written as:

$$y_i = \left(\alpha_i + \sum_j \alpha_{ij} \overline{q}_j \right) \quad i = 1,...,n-1$$

(4) $\qquad x_h = -\left(\beta_h + \sum_j \beta_{hj} \overline{q}_j \right) h = 1,...,m$

$$s_i = \left(\gamma_i + \sum_j \gamma_{ij} \overline{q}_j \right) \quad i = 1,...,n_p$$

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 $\overline{\mathbf{q}}$, 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 $\gamma_{ij} = \gamma_{ji}$.

3. 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-year period, for a total of 34,140 observations[2].

---

[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

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, …), 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 (maize, durum wheat, other cereals, oilseeds, 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) and two fixed inputs (total land and an 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 prices[3]. 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 expectations[4].

4. Econometric techniques

4.1 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

---

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.

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

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

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 $f_i(.)$ 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 $f_i(.)$; $\Phi(.)$ and $\Theta(.)$ 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/altitudes[5]. Thus, in each probit model we estimated 11 parameters, including a constant term.

4.2 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 Chaaban and Thomas (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 disturbance[6]. In order to choose the right specification between RE and FE, one may rely on the Hausman test, which is based on Generalised Least Square estimation (Baltagi, 2005).

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 since we like to carry out the corresponding Hausman test for choosing between two-way FE and two-way RE, we have decided to build our own GLS estimator.

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

---

[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 for example 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 (Moro and Sckokai, 1999; Sckokai and Moro, 2006; Serra et al, 2005; Serra et al., 2006).

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.

4.2.1 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 $i=1,...,F$, and by T periods, indexed by $t=1,...,T$. 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 $\boldsymbol{D_t}$ be the $F_t \times F$ matrix obtained from I_F by omitting the rows corresponding to farms not observed in period t. With $\underset{n \times F}{\boldsymbol{\Delta_1}} = \left(\boldsymbol{D'_1},..., \boldsymbol{D'_T}\right)'$ and $\underset{n \times T}{\boldsymbol{\Delta_2}} = diag\left(\boldsymbol{D_t \iota_F}\right)$ we can define $\underset{F \times F}{\boldsymbol{\Delta_F}} \equiv \boldsymbol{\Delta'_1 \Delta_1}$, $\underset{T \times T}{\boldsymbol{\Delta_T}} \equiv \boldsymbol{\Delta'_2 \Delta_2}$ and $\underset{T \times F}{\boldsymbol{\Delta_{TF}}} \equiv \boldsymbol{\Delta'_2 \Delta_1}$ and we can consider $\boldsymbol{\Delta} \equiv \underset{n \times F \quad n \times T}{\left(\boldsymbol{\Delta_1}, \boldsymbol{\Delta_2}\right)}$ which gives the dummy-variable structure for the unbalanced panel model.

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

$$(6) \overline{\boldsymbol{\Delta}} \equiv \boldsymbol{\Delta_2} - \boldsymbol{\Delta_1 \Delta_F^{-1} \Delta'_{TF}} = \left(\boldsymbol{I_F} - \boldsymbol{\Delta_1 \Delta_F^{-1} \Delta'_1}\right) \cdot \boldsymbol{\Delta_2} = \left(\boldsymbol{I_F} - \boldsymbol{P_{[\Delta_1]}}\right) \cdot \boldsymbol{\Delta_2} = \boldsymbol{Q_{[\Delta_1]} \Delta_2}$$

$$(7) \boldsymbol{Q} \equiv \boldsymbol{\Delta_T} - \boldsymbol{\Delta_{TF} \Delta_F^{-1} \Delta'_{TF}} = \boldsymbol{\Delta'_2} \cdot \left(\boldsymbol{\Delta_2} - \boldsymbol{\Delta_1 \Delta_F^{-1} \Delta'_{TF}}\right) = \boldsymbol{\Delta'_2 \overline{\Delta}} = \boldsymbol{\Delta'_2 Q_{[\Delta_1]} \Delta_2}$$

and hence the projection matrix onto the null-space of $\boldsymbol{\Delta}$ is:

$$(8) \boldsymbol{Q_{[\Delta]}} = \left(\boldsymbol{I_F} - \boldsymbol{\Delta_1 \Delta_F^{-1} \Delta'_1}\right) - \overline{\boldsymbol{\Delta}} \boldsymbol{Q^- \overline{\Delta}'} = \boldsymbol{Q_{[\Delta_1]}} - \boldsymbol{Q_{[\Delta_1]} \Delta_2 Q^- \Delta'_2 Q_{[\Delta_1]}}$$

and therefore the within estimator is:

$$(9) \boldsymbol{\beta^{WT}} = \left(\boldsymbol{X'Q_{[\Delta]}X}\right)^{-1} \left(\boldsymbol{X'Q_{[\Delta]}y}\right)$$

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

$$(10) \boldsymbol{\Omega} = \boldsymbol{E}\left(\boldsymbol{uu'}\right) = \sigma_\varepsilon^2 \boldsymbol{I_n} + \sigma_\mu^2 \boldsymbol{\Delta_1 \Delta'_1} + \sigma_\nu^2 \boldsymbol{\Delta_2 \Delta'_2}$$

With $\widetilde{\boldsymbol{\Delta}}_F = \boldsymbol{\Delta_F} + \dfrac{\sigma_\varepsilon^2}{\sigma_\mu^2} \boldsymbol{I_F}$ and $\widetilde{\boldsymbol{\Delta}}_T = \boldsymbol{\Delta_T} + \dfrac{\sigma_\varepsilon^2}{\sigma_\nu^2} \boldsymbol{I_T}$, by defining

$$\widetilde{\boldsymbol{Q}} = \widetilde{\boldsymbol{\Delta}}_T - \boldsymbol{\Delta_{TF}} \widetilde{\boldsymbol{\Delta}}_F^{-1} \boldsymbol{\Delta'_{TF}}$$

$$\widetilde{\boldsymbol{Q}}_{[\Delta_1]} = \boldsymbol{I_F} - \boldsymbol{\Delta_1} \widetilde{\boldsymbol{\Delta}}_F^{-1} \boldsymbol{\Delta'_1}$$

Wansbeek and Kapteyn (1989) show that

$$(11) \boldsymbol{\Omega^{-1}} = \frac{1}{\sigma_\varepsilon^2} \left(\widetilde{\boldsymbol{Q}}_{[\Delta_1]} - \widetilde{\boldsymbol{Q}}_{[\Delta_1]} \boldsymbol{\Delta_2} \widetilde{\boldsymbol{Q}}^- \boldsymbol{\Delta'_2} \widetilde{\boldsymbol{Q}}_{[\Delta_1]}\right)$$

and then the *GLS* estimator is

$$(12)\ \beta^{GLS} = \left(X'\Omega^{-1}X\right)^{-1}\left(X'\Omega^{-1}y\right)$$

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

$$q_n = f'Q_{[\Delta]}f$$

$$(13)\qquad q_F = f'\Delta_2\Delta_T^{-1}\Delta_2'f$$

$$q_T = f'\Delta_1\Delta_F^{-1}\Delta_1'f$$

to their expected values

$$E(q_n) = \left(n - T - F + 1 - k\right)\cdot\sigma_\varepsilon^2$$

$$(14)\qquad E(q_F) = \left(T + k_F - k_0 - 1\right)\cdot\sigma_\varepsilon^2 + T\cdot\sigma_\mu^2 + n\cdot\sigma_\nu^2 - \left(\lambda_1\sigma_\mu^2 + \lambda_1\sigma_\nu^2\right)$$

$$E(q_T) = \left(F + k_T - k_0 - 1\right)\cdot\sigma_\varepsilon^2 + n\cdot\sigma_\mu^2 + F\cdot\sigma_\nu^2 - \left(\lambda_1\sigma_\mu^2 + \lambda_1\sigma_\nu^2\right)$$

with $\qquad k_F = \mathrm{tr}\left\{\left(X'Q_{[\Delta]}X\right)^{-1}X'\Delta_2\Delta_T^{-1}\Delta_2'X\right\}$, $\qquad k_T = \mathrm{tr}\left\{\left(X'Q_{[\Delta]}X\right)^{-1}X'\Delta_1\Delta_F^{-1}\Delta_1'X\right\}$,

$$k_0 = \frac{\iota_n'X\left(X'Q_{[\Delta]}X\right)^{-1}X'\iota_n}{n},\ \lambda_1 = \frac{\iota_n'\Delta_1\Delta_1'\iota_n}{n}\ \text{and}\ \lambda_2 = \frac{\iota_n'\Delta_2\Delta_2'\iota_n}{n}$$

4.2.2 Two-Way ECM and SUR estimation

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, since each of our dependent variables (output supplies, variable input demands and land allocations) is regressed over the same set of explanatory variables (output and input prices, area payments and quasi-fixed inputs). Once again, standard econometric softwares do not provide automatic commands to estimate two-way ECM 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 ECM SUR*. Based on this framework, we have derived the corresponding estimator for the *Two-Way ECM SUR*.

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

$$(15) \qquad \underset{M \times 1}{\boldsymbol{y}_{it}} = \underset{M \times k_{sur}}{\boldsymbol{X}_{it}} \underset{k_{sur} \times 1}{\boldsymbol{\beta}} + \underset{M \times 1}{\boldsymbol{\mu}_i} + \underset{M \times 1}{\boldsymbol{\nu}_t} + \underset{M \times 1}{\boldsymbol{\varepsilon}_{it}} = \underset{M \times k_{sur}}{\boldsymbol{X}_{it}} \underset{k_{sur} \times 1}{\boldsymbol{\beta}} + \underset{M \times 1}{\boldsymbol{u}_{it}}$$

where $\boldsymbol{X}_{it} = \mathrm{diag}(\boldsymbol{x}_{1it},...,\boldsymbol{x}_{Mit})$ and $\boldsymbol{\beta} = \mathrm{diag}(\boldsymbol{\beta}_1',...,\boldsymbol{\beta}_M')'$. With $\boldsymbol{\mu}_i = (\mu_{1i},...,\mu_{Mi})'$, $\boldsymbol{\nu}_t = (\nu_{1t},...,\nu_{Mt})'$ and $\boldsymbol{\varepsilon}_{it} = (\varepsilon_{1it},...,\varepsilon_{Mit})'$, we assume

$$(16) \qquad
\begin{aligned}
\mathrm{E}(\mu_{mi},\mu_{ji'}) &= \sigma^2_{\mu_{mj}} & i = i' \\
&= 0 & i \neq i' \\
\mathrm{E}(\nu_{mt},\nu_{jt'}) &= \sigma^2_{\nu_{mj}} & t = t' \\
&= 0 & t \neq t' \\
\mathrm{E}(\varepsilon_{mit},\varepsilon_{ji't'}) &= \sigma^2_{\varepsilon_{mj}} & i = i' \ and \ t = t' \\
&= 0 & i \neq i' \ or \ t \neq t'
\end{aligned}$$

Hence $\boldsymbol{\mu}_i$, $\boldsymbol{\nu}_t$ and $\boldsymbol{\varepsilon}_{it}$ have zero expectations and covariance matrices $\boldsymbol{\Sigma}_\mu$, $\boldsymbol{\Sigma}_\nu$ and $\boldsymbol{\Sigma}_\varepsilon$. It follows that $E(\boldsymbol{u}_{it}\boldsymbol{u}_{i't'}') = \delta_{ii'}\boldsymbol{\Sigma}_\mu + \delta_{tt'}\boldsymbol{\Sigma}_\nu + \delta_{ii'}\delta_{tt'}\boldsymbol{\Sigma}_\varepsilon$ with $\delta_{ii'} = 1$ for $i = i'$ and $\delta_{ii'} = 0$ for $i \neq i'$, $\delta_{tt'} = 1$ for $t = t'$ and $\delta_{tt'} = 0$ for $t \neq t'$.

Let us consider $\boldsymbol{y}_{i(p)} = (\boldsymbol{y}_{i1}',...,\boldsymbol{y}_{ip}')'$, $\boldsymbol{X}_{i(p)} = (\boldsymbol{X}_{i1}',...,\boldsymbol{X}_{ip}')'$ and $\boldsymbol{u}_{i(p)} = (\boldsymbol{u}_{i1}',...,\boldsymbol{u}_{ip}')'$ for $i \in I_p$ and $p = 2,...,P$ (and then for $i = 1,...,C_2, C_2 + 1,...,C_3,...,C_{P-1} + 1,...,C_P$ with $C_P = F$).

We define the matrix $\underset{pM \times TM}{\boldsymbol{\Delta}_{i(p)}}$ 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 $\underset{pM \times TM}{\boldsymbol{\Delta}_{i(p)}} = \begin{bmatrix} \boldsymbol{0} & \boldsymbol{\iota}_M & \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{\iota}_M \end{bmatrix}$. Hence with $\underset{TM \times 1}{\boldsymbol{\nu}} = (\boldsymbol{\nu}_1',...,\boldsymbol{\nu}_T')'$ for the farm $i \in I_p$ we can define $\underset{pM \times 1}{\boldsymbol{\nu}_{i(p)}} = \underset{pM \times TM}{\boldsymbol{\Delta}_{i(p)}} \underset{TM \times 1}{\boldsymbol{\nu}}$ and we can write the model

$$\underset{pM \times 1}{\boldsymbol{y}_{i(p)}} = \underset{pM \times k_{sur}}{\boldsymbol{X}_{i(p)}} \underset{k_{sur} \times 1}{\boldsymbol{\beta}} + \left(\underset{p \times 1}{\boldsymbol{\iota}_p} \otimes \underset{M \times 1}{\boldsymbol{\mu}_i} \right) + \underset{pM \times 1}{\boldsymbol{\nu}_{i(p)}} + \underset{pM \times 1}{\boldsymbol{\varepsilon}_{i(p)}} = \underset{pM \times k_{sur}}{\boldsymbol{X}_{i(p)}} \underset{k_{sur} \times 1}{\boldsymbol{\beta}} + \underset{pM \times 1}{\boldsymbol{u}_{i(p)}}$$

$$(17) \qquad
\begin{aligned}
\boldsymbol{\Omega}_p &= & \boldsymbol{I}_p \otimes (\boldsymbol{\Sigma}_\varepsilon + \boldsymbol{\Sigma}_\nu) + \boldsymbol{J}_p \otimes \boldsymbol{\Sigma}_\mu & = \\
&= & \boldsymbol{E}_p \otimes (\boldsymbol{\Sigma}_\varepsilon + \boldsymbol{\Sigma}_\nu) + \bar{\boldsymbol{J}}_p \otimes (\boldsymbol{\Sigma}_\varepsilon + \boldsymbol{\Sigma}_\nu) + \boldsymbol{J}_p \otimes \boldsymbol{\Sigma}_\mu & = \\
&= & \boldsymbol{E}_p \otimes (\boldsymbol{\Sigma}_\varepsilon + \boldsymbol{\Sigma}_\nu) + \bar{\boldsymbol{J}}_p \otimes (\boldsymbol{\Sigma}_\varepsilon + \boldsymbol{\Sigma}_\nu) + \bar{\boldsymbol{J}}_p \otimes p\boldsymbol{\Sigma}_\mu & = \\
&= & \boldsymbol{E}_p \otimes (\boldsymbol{\Sigma}_\varepsilon + \boldsymbol{\Sigma}_\nu) + \bar{\boldsymbol{J}}_p \otimes (\boldsymbol{\Sigma}_\varepsilon + \boldsymbol{\Sigma}_\nu + p\boldsymbol{\Sigma}_\mu)
\end{aligned}$$

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

$$(18) \qquad \boldsymbol{\Omega}_p^{-1} = \boldsymbol{E}_p \otimes (\boldsymbol{\Sigma}_\varepsilon + \boldsymbol{\Sigma}_\nu)^{-1} + \bar{\boldsymbol{J}}_p \otimes (\boldsymbol{\Sigma}_\varepsilon + \boldsymbol{\Sigma}_\nu + p\boldsymbol{\Sigma}_\mu)^{-1}$$

Then we can consider the *GLS* problem for $\boldsymbol{\beta}$ when $\boldsymbol{\Sigma}_\mu$, $\boldsymbol{\Sigma}_\nu$ and $\boldsymbol{\Sigma}_\varepsilon$ are known, i.e. the problem of minimizing

$$\sum_{p=2}^{P}\sum_{i\in I_p}\underset{1\times pM}{u'_{i(p)}}\ \underset{pM\times pM}{\Omega_p^{-1}}\ \underset{pM\times 1}{u_{i(p)}}=\sum_{p=2}^{P}\sum_{i\in I_p}\underset{1\times pM}{u'_{i(p)}}\Big[E_p\otimes\underset{pM\times pM}{\big(\Sigma_{\varepsilon}+\Sigma_{\nu}\big)^{-1}}\Big]\underset{pM\times 1}{u_{i(p)}}+\sum_{p=2}^{P}\sum_{i\in I_p}\underset{1\times pM}{u'_{i(p)}}\Big[\bar{J}_p\otimes\underset{pM\times pM}{\big(\Sigma_{\varepsilon}+\Sigma_{\nu}+p\Sigma_{\mu}\big)^{-1}}\Big]\underset{pM\times 1}{u_{i(p)}}$$

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

$$\beta_p^{GLS}=\left[\sum_{i\in I_p}\underset{1\times pM}{X'_{i(p)}}\ \underset{pM\times pM}{\Omega_p^{-1}}\ \underset{pM\times 1}{X_{i(p)}}\right]^{-1}\times\left[\sum_{i\in I_p}\underset{1\times pM}{X'_{i(p)}}\ \underset{pM\times pM}{\Omega_p^{-1}}\ \underset{pM\times 1}{y_{i(p)}}\right]=$$

$$=\left[\sum_{i\in I_p}\underset{1\times pM}{X'_{i(p)}}\Big[E_p\otimes\underset{pM\times pM}{\big(\Sigma_{\varepsilon}+\Sigma_{\nu}\big)^{-1}}\Big]\underset{pM\times 1}{X_{i(p)}}+\sum_{i\in I_p}\underset{1\times pM}{X'_{i(p)}}\Big[\bar{J}_p\otimes\underset{pM\times pM}{\big(\Sigma_{\varepsilon}+\Sigma_{\nu}+p\Sigma_{\mu}\big)^{-1}}\Big]\underset{pM\times 1}{X_{i(p)}}\right]^{-1}\times$$

$$\times\left[\sum_{i\in I_p}\underset{1\times pM}{X'_{i(p)}}\Big[E_p\otimes\underset{pM\times pM}{\big(\Sigma_{\varepsilon}+\Sigma_{\nu}\big)^{-1}}\Big]\underset{pM\times 1}{y_{i(p)}}+\sum_{i\in I_p}\underset{1\times pM}{X'_{i(p)}}\Big[\bar{J}_p\otimes\underset{pM\times pM}{\big(\Sigma_{\varepsilon}+\Sigma_{\nu}+p\Sigma_{\mu}\big)^{-1}}\Big]\underset{pM\times 1}{y_{i(p)}}\right]$$

while the full *GLS* estimator is

$$\beta^{GLS}=\left[\sum_{p=2}^{P}\sum_{i\in I_p}\underset{1\times pM}{X'_{i(p)}}\ \underset{pM\times pM}{\Omega_p^{-1}}\ \underset{pM\times 1}{X_{i(p)}}\right]^{-1}\times\left[\sum_{p=2}^{P}\sum_{i\in I_p}\underset{1\times pM}{X'_{i(p)}}\ \underset{pM\times pM}{\Omega_p^{-1}}\ \underset{pM\times 1}{y_{i(p)}}\right]=$$

$$=\left[\sum_{p=2}^{P}\sum_{i\in I_p}\underset{1\times pM}{X'_{i(p)}}\Big[E_p\otimes\underset{pM\times pM}{\big(\Sigma_{\varepsilon}+\Sigma_{\nu}\big)^{-1}}\Big]\underset{pM\times 1}{X_{i(p)}}+\sum_{p=2}^{P}\sum_{i\in I_p}\underset{1\times pM}{X'_{i(p)}}\Big[\bar{J}_p\otimes\underset{pM\times pM}{\big(\Sigma_{\varepsilon}+\Sigma_{\nu}+p\Sigma_{\mu}\big)^{-1}}\Big]\underset{pM\times 1}{X_{i(p)}}\right]^{-1}\times$$

$$\times\left[\sum_{p=2}^{P}\sum_{i\in I_p}\underset{1\times pM}{X'_{i(p)}}\Big[E_p\otimes\underset{pM\times pM}{\big(\Sigma_{\varepsilon}+\Sigma_{\nu}\big)^{-1}}\Big]\underset{pM\times 1}{y_{i(p)}}+\sum_{p=2}^{P}\sum_{i\in I_p}\underset{1\times pM}{X'_{i(p)}}\Big[\bar{J}_p\otimes\underset{pM\times pM}{\big(\Sigma_{\varepsilon}+\Sigma_{\nu}+p\Sigma_{\mu}\big)^{-1}}\Big]\underset{pM\times 1}{y_{i(p)}}\right]$$

We can estimate the covariance matrices Σ_{μ}, Σ_{ν} 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 Wansbeed and Kapteyn (1989) – corrected for the *SUR.*.

The first method considers the *FE residuals* $\underset{M\times 1}{e_{it}}\equiv y_{it}-X\beta^{WT}$ for the farm *i* in period *t*. If we define $f_{it}=e_{it}-\bar{e}$, the $M\times M$ matrices of within farms, between farms and between times (co)variations in the *f*'s of the different equations are

$$W_f=\sum_{i=1}^{F}\sum_{t=1}^{T_i}\big(f_{it}-\bar{f}_i-\bar{f}_t\big)\big(\bar{f}_{it}-\bar{f}_i-\bar{f}_t\big)'$$

$$B_f^C=\sum_{i=1}^{F}T_i\big(\bar{f}_i-\bar{f}\big)\big(\bar{f}_i-\bar{f}\big)'$$

$$B_f^T=\sum_{i=1}^{T}F_t\big(\bar{f}_t-\bar{f}\big)\big(\bar{f}_t-\bar{f}\big)'$$

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

$$W_f=E\big(W_{\varepsilon}\big)$$

$$B_f^C=E\big(B_{\mu}^C\big)+E\big(B_{\varepsilon}^C\big)$$

$$B_f^T=E\big(B_{\nu}^T\big)+E\big(B_{\varepsilon}^T\big)$$

where

$$W_{\varepsilon} = \sum_{i=1}^{F}\sum_{t=1}^{T_i} \varepsilon_{it}\varepsilon_{it}^{'} - \sum_{i=1}^{F} T_i \bar{\varepsilon}_i \bar{\varepsilon}_i^{'} - \sum_{t=1}^{T} F_t \bar{\varepsilon}_t \bar{\varepsilon}_t^{'}$$

$$B_{\mu}^{C} = \sum_{i=1}^{F} T_i \mu_i \mu_i^{'} - \sum_{i=1}^{F} T_i \overline{\mu\mu}' = \sum_{i=1}^{F} T_i \mu_i \mu_i^{'} - n\overline{\mu\mu}'$$

$$B_{\varepsilon}^{C} = \sum_{i=1}^{F} T_i \bar{\varepsilon}_i \bar{\varepsilon}_i^{'} - \sum_{i=1}^{F} T_i \overline{\varepsilon\varepsilon}' = \sum_{i=1}^{F} T_i \bar{\varepsilon}_i \bar{\varepsilon}_i^{'} - n\overline{\varepsilon\varepsilon}'$$

$$B_{\nu}^{T} = \sum_{t=1}^{T} F_t \nu_t \nu_t^{'} - \sum_{t=1}^{T} F_t \overline{\nu\nu}' = \sum_{t=1}^{T} F_t \nu_t \nu_t^{'} - n\overline{\nu\nu}'$$

$$B_{\varepsilon}^{T} = \sum_{t=1}^{T} F_t \bar{\varepsilon}_t \bar{\varepsilon}_t^{'} - \sum_{t=1}^{T} F_t \overline{\varepsilon\varepsilon}' = \sum_{t=1}^{T} F_t \bar{\varepsilon}_t \bar{\varepsilon}_t^{'} - n\overline{\varepsilon\varepsilon}'$$

Since $\quad E(u_{it}u_{i't'}^{'}) = \delta_{ii'}\Sigma_{\mu} + \delta_{tt'}\Sigma_{\nu} + \delta_{ii'}\delta_{tt'}\Sigma_{\varepsilon}$, $\qquad E(\mu_i\mu_{i'}^{'}) = \delta_{ii'}\Sigma_{\mu}$, $\qquad E(\nu_t\nu_{t'}^{'}) = \delta_{tt'}\Sigma_{\nu}$ $\qquad$ and

$E(\varepsilon_t\varepsilon_{t'}^{'}) = \delta_{ii'}\delta_{tt'}\Sigma_{\varepsilon}$ it follows that

$$E(W_{\varepsilon}) = (n - F - T) \cdot \Sigma_{\varepsilon}$$

$$E(B_{\mu}^{C}) = \left(n - \frac{\sum_{i=1}^{F} T_i^2}{n} \right) \cdot \Sigma_{\mu}$$

$$E(B_{\varepsilon}^{C}) = (F - 1) \cdot \Sigma_{\varepsilon}$$

$$E(B_{\nu}^{T}) = \left(n - \frac{\sum_{i=1}^{T} F_t^2}{n} \right) \cdot \Sigma_{\nu}$$

$$E(B_{\varepsilon}^{T}) = (T - 1) \cdot \Sigma_{\varepsilon}$$

Hence we can conclude that

$$\hat{\Sigma}_{\varepsilon} = \frac{W_f}{n - F - T}$$

$$\hat{\Sigma}_{\mu} = \frac{B_f^{C} - (F - 1) \cdot \hat{\Sigma}_{\varepsilon}}{n - \dfrac{\sum_{i=1}^{F} T_i^2}{n}}$$

$$\hat{\Sigma}_{\nu} = \frac{B_f^{T} - (T - 1) \cdot \hat{\Sigma}_{\varepsilon}}{n - \dfrac{\sum_{i=1}^{T} F_t^2}{n}}$$

The second method considers the *FE residuals* $e_m \equiv y_m - X\beta_m^{WT}$ for the equation $m=1,..., M$.

If we define $f_m = E_n \cdot e_m = e_m - \bar{e}_m$ we can obtain *QUE's* for $\sigma_{\varepsilon_{mj}}^2$, $\sigma_{\mu_{mj}}^2$ and $\sigma_{\nu_{mj}}^2$ by equating

$$q_{n_{mj}} \equiv f_j^{'} Q_{[\Delta]} f_m$$

(19) $q_{F_{mj}} \equiv \boldsymbol{f}_j' \boldsymbol{\Delta}_2 \boldsymbol{\Delta}_T^{-1} \boldsymbol{\Delta}_2' \boldsymbol{f}_m$

$\qquad q_{T_{mj}} \equiv \boldsymbol{f}_j' \boldsymbol{\Delta}_1 \boldsymbol{\Delta}_F^{-1} \boldsymbol{\Delta}_1' \boldsymbol{f}_m$

to their expected values

$$E\big(q_{n_{mj}}\big) = \big(n - T - F + 2 - k_m - k_j + k_{mj}\big) \cdot \sigma^2_{\varepsilon_{mj}}$$

(20) $E\big(q_{F_{mj}}\big) = \big(T + k_{F_{mj}} - k_{0_{mj}} - 1\big) \cdot \sigma^2_{\varepsilon_{mj}} + T \cdot \sigma^2_{\mu_{mj}} + n \cdot \sigma^2_{\nu_{mj}} - \big(\lambda_1 \sigma^2_{\mu_{mj}} + \lambda_1 \sigma^2_{\nu_{mj}}\big)$

$\qquad E\big(q_{T_{mj}}\big) = \big(F + k_{T_{mj}} - k_{0_{mj}} - 1\big) \cdot \sigma^2_{\varepsilon_{mj}} + n \cdot \sigma^2_{\mu_{mj}} + F \cdot \sigma^2_{\nu_{mj}} - \big(\lambda_1 \sigma^2_{\mu_{mj}} + \lambda_1 \sigma^2_{\nu_{mj}}\big)$

with

$$k_{mj} \equiv \operatorname{tr}\Big\{\big(X_m' Q_{[\Delta]} X_m\big)^{-1}\big(X_m' Q_{[\Delta]} X_j\big)\big(X_j' Q_{[\Delta]} X_j\big)^{-1}\big(X_j' Q_{[\Delta]} X_m\big)\Big\}$$

$$k_{F_{mj}} \equiv \operatorname{tr}\Big\{\big(X_m' Q_{[\Delta]} X_j\big)^{-1}\big(X_j' Q_{[\Delta]} X_m\big)\big(X_m' Q_{[\Delta]} X_m\big)^{-1} X_j' \Delta_2 \Delta_T^{-1} \Delta_2' X_m\Big\}$$

$$k_{T_{mj}} \equiv \operatorname{tr}\Big\{\big(X_m' Q_{[\Delta]} X_j\big)^{-1}\big(X_j' Q_{[\Delta]} X_m\big)\big(X_m' Q_{[\Delta]} X_m\big)^{-1} X_j' \Delta_1 \Delta_F^{-1} \Delta_1' X_m\Big\}$$

$$k_{0_{mj}} = \frac{\iota_n' X_m \big(X_m' Q_{[\Delta]} X_m\big)^{-1} X_m' Q_{[\Delta]} X_j \big(X_j' Q_{[\Delta]} X_j\big)^{-1} X_j' \iota_n}{n}$$

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 $\sigma^2_{\varepsilon_{mj}} = \sigma^2_{\varepsilon_{jm}}$, $\sigma^2_{\mu_{mj}} = \sigma^2_{\mu_{jm}}$ and $\sigma^2_{\nu_{mj}} = \sigma^2_{\nu_{jm}}$.

5. Results

For space reasons, we cannot report all the estimated parameters. However, it is important to note that, carrying out the Hausman tests for all the equations of the system, we always get a χ^2 value statistically significant at the 1% level, which means that we reject the null hypothesis of zero random effects. This implies that the RE specification is always preferred to the FE specification.

For illustrative purposes, in Table 1 we report the estimated parameters for the durum wheat supply equation, one of the most important arable crop in Italy. 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 significance[7].

For example, the own price-response of durum wheat (the P_2 row in Table 1) is significant only in the 1-way FE and RE models and in the 2-way SUR system, while in other models is not statistically significant. Moreover, among significant parameters, we observe quite a strong variability, since the 2-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 A_2 row in the same table). Here all models provide positive and significant parameters, but their absolute value is strongly different among models, with a different ranking with respect to the own-price parameter discussed above (i.e. here the highest value is obtained through the 1-way RE model, while the 2-way SUR provides a value which is much closer to the FE parameters).

---

[7] In Tables 1 and 2, we adopt the following convention: Q= Supply, P= Price, S= land allocation, A= area payment, 1= Other Cereals, 2= Durum wheat, 3= Oilseeds, 4= Maize, 5= Other output; 6= Crop input, S= Total Land, E= Set-aside Percentage, Z=Quasi fixed inputs.

In Table 2, 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 for both maize supply and land allocation (rows Q_4 and S_4 in the Table) the own payment elasticity is much higher in the case of the two-way SUR system as compared to all the other models. For the corresponding price elasticity, the two-way SUR system provides a negative and significant elasticity (which is of course counterintuitive!) , while a number of other models estimates a positive response to price.

6. 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. 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), while the choice between the alternative Fixed Effect (FE) and Random Effect (RE) specifications has been carried out through the appropriate Hausman tests. 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 elasticites estimated in our model.

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

Q2	FE 1 WAY	FE 2 WAY	RE 1 WAY (ML)	RE 1 WAY (GLS)	RE 2 WAY (ML)	RE 2 WAY (GLS)	SUR 2 WAY WB (GLS)	SUR 2 WAY QUE (GLS)
P1	-1.3836 *** (0.1441)	-1.3270 *** (0.1447)	-1.2476 *** (0.1318)	-1.2426 *** (0.1300)	-1.1916 *** (0.1319)	-1.2118 *** (0.1341)	-1.4894 *** (0.1252)	-1.9022 *** (0.1232)
P2	0.7387 ** (0.3162)	-0.3754 (0.3816)	0.6598 ** (0.3078)	0.6600 ** (0.3048)	-0.2403 (0.3668)	-0.2437 (0.3651)	0.9090 *** (0.3009)	0.9343 *** (0.2962)
P3	-0.6987 *** (0.1676)	-0.7061 *** (0.1685)	-0.3569 ** (0.1570)	-0.3362 ** (0.1550)	-0.3529 ** (0.1573)	-0.4291 *** (0.1591)	-0.6952 *** (0.1501)	-0.8854 *** (0.1474)
P4	0.5820 *** (0.1807)	0.9533 *** (0.1868)	0.4411 ** (0.1707)	0.4366 ** (0.1687)	0.7726 *** (0.1756)	0.7858 *** (0.1769)	0.2802 * (0.1644)	0.4750 *** (0.1627)
P5	-0.1466 (0.1329)	-0.3742 ** (0.1678)	-0.4294 *** (0.1279)	-0.4464 *** (0.1265)	-0.6780 *** (0.1584)	-0.6008 *** (0.1579)	-0.7750 *** (0.1243)	-0.5041 *** (0.1225)
P6	-4.4451 *** (1.2687)	-2.4421 * (1.3965)	-7.4735 *** (1.2322)	-7.6837 *** (1.2187)	-5.6935 *** (1.3493)	-4.9683 *** (1.3469)	-7.3669 *** (1.2028)	-5.8880 *** (1.1873)
A1	0.0434 (0.1952)	-0.0272 (0.2028)	-0.2052 (0.1742)	-0.2085 (0.1717)	-0.2899 (0.1793)	-0.2749 (0.1830)	0.6242 *** (0.1637)	0.4295 *** (0.1597)
A2	0.4052 *** (0.0934)	0.4515 *** (0.0942)	0.9928 *** (0.0872)	1.0263 *** (0.0856)	1.0248 *** (0.0875)	0.9040 *** (0.0883)	0.9574 *** (0.0826)	0.5926 *** (0.0816)
A3	-0.0536 (0.0342)	0.0845 ** (0.0402)	-0.0975 *** (0.0334)	-0.0996 *** (0.0330)	0.0434 *** (0.0390)	0.0459 (0.0385)	-0.0789 ** (0.0326)	-0.0384 (0.0322)
A4	-0.0491 (0.0569)	-0.0499 (0.0604)	-0.0531 (0.0542)	-0.0537 (0.0536)	-0.0747 (0.0571)	-0.0679 (0.0574)	-0.0329 (0.0526)	0.0673 (0.0520)
S	0.5398 *** (0.0125)	0.5369 *** (0.0125)	0.4203 *** (0.0076)	0.4174 *** (0.0073)	0.4190 (0.0076)	0.4312 *** (0.0081)	0.4947 *** (0.0073)	0.4121 *** (0.0072)
Z	0.0011 *** (0.0003)	0.0011 *** (0.0003)	0.0016 *** (0.0002)	0.0017 *** (0.0002)	0.0016 *** (0.0002)	0.0013 *** (0.0003)	0.0017 *** (0.0002)	0.0030 *** (0.0002)
E	-37.0379 *** (2.4832)	-37.3892 *** (2.4926)	-24.4107 *** (2.3250)	-23.6270 *** (2.2858)	-24.6105 *** (2.3273)	-27.4723 *** (2.3467)	-31.8690 *** (2.2148)	-35.7399 *** (2.1843)

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

Table 2: Estimated own-price and own-payment elasticities under different panel-data techniques

		FE 1 WAY	FE 2 WAY	RE 1 WAY (ML)	RE 1 WAY (GLS)	RE 2 WAY (ML)	RE 2 WAY (GLS)	SUR (WB) 2 WAY	SUR (QUE) 2 WAY
Q1	**price**	0.5797 ***	0.5806 ***	0.5765 ***	0.5774 ***	0.5767 ***	0.5753 ***	0.6011 ***	0.6282 ***
	payment	-0.0624	-0.0773	0.0649	0.0942 *	0.0524	0.0494	-0.0813	-0.1317 **
Q2	**price**	0.1206 **	-0.0613	0.1077 **	0.1077 **	-0.0392	-0.0398	0.1484 ***	0.1525 ***
	payment	0.1571 ***	0.1751 ***	0.3850 ***	0.3980 ***	0.3974 ***	0.3506 ***	0.3713 ***	0.2298 ***
Q3	**price**	0.7875 ***	0.7712 ***	0.8607 ***	0.8775 ***	0.8459 ***	0.8633 ***	0.8785 ***	0.8746 ***
	payment	-0.0027	0.0161	0.0486	0.0561	0.0349	0.0777	0.0982 *	0.1153 **
Q4	**price**	0.0280	0.0650	0.0650	0.0650	0.0848 *	0.0766 *	0.0496	-0.1655 ***
	payment	0.0599 **	-0.0259	0.0886 ***	0.0886 ***	0.0219	0.0053	0.0642 **	0.3206 ***
S1	**price**	0.5868	0.5861	0.5995 ***	0.6063 ***	0.5985 ***	0.5951 ***	0.6428 ***	0.6573
	payment	-0.0807	-0.0952	-0.0141	-0.0265	-0.0260	-0.0303 ***	-0.1909	-0.2042
S2	**price**	-0.0530 ***	-0.0765 ***	-0.0622 ***	-0.0619 ***	-0.0825 ***	-0.0829 ***	-0.0320 ***	-0.0259 ***
	payment	0.0434 *	0.0543 *	0.1727	0.1877	0.1808	0.1576	0.1623 ***	0.0327 **
S3	**price**	0.7910	0.7714	0.8689	0.8836	0.8509 *	0.8672 *	0.8961	0.8983 ***
	payment	0.0923	0.0358 *	0.0814 ***	0.0773 ***	0.0169 ***	0.0580 ***	0.0743 ***	0.0930
S4	**price**	0.0657 ***	0.0604 ***	0.0983 ***	0.0989 ***	0.0870 ***	0.0761 ***	0.0632 ***	-0.2073 ***
	payment	0.0805 **	-0.0126	0.1183 **	0.1192 **	0.0430	0.0242	0.1127 *	0.3185 **

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

References

Baltagi, B.H., Song, S.H. and Jung, B.C. (2001). The Unbalanced Nested Error Component Regression Model. *Journal of Econometrics* 1061: 357-381.

Baltagi, B.H. (2005). *Econometric Analysis of Panel Data (3rd edition)*. Chichester: Wiley and Sons.

Biørn, E. (2004). Regression Systems for Unbalanced Panel Data: a Stepwise Maximum Likelihood Procedure. *Journal of Econometrics* 122: 281-291.

Boumahdi, R., Chaaban, J. and 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.

Chavas, J.P. and Holt, M.T. (1990). Acreage Decision Under Risk: The Case of Corn and Soybeans. *American Journal of Agricultural Economics* 72: 529-538.

Davis, P. (2002). Estimating Multi-way Error Components Models with Unbalanced Data Structures. *Journal of Econometrics* 106: 67-95.

Goodwin, B.K and Mishra A.K., (2006). Are 'Decoupled' Farm Program Payments Really Decoupled? An Empirical Evaluation. *American Journal of Agricultural Economics* 88: 73-89.

Greene, W. (2003). *Econometric Analysis, 5th edition*. New York: Prentice Hall.

Moro, D. and 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.

Oude Lansink, A. and Peerlings, J. (1996). Modelling the New EU Cereals Regime in the Netherlands. *European Review of Agricultural Economics* 23: 161-178.

Oude Lansink, A. (1999). Area Allocation Under Price Uncertainty on Dutch Arable Farms. *Journal of Agricultural Economics* 50:, 93-105.

Sckokai, P. and 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.

Sckokai, P. and Moro, D. (2006). Modeling the Reforms of the Common Agricultural Policy for Arable Crops under Uncertainty. *American Journal of Agricultural Economics* 88: 43-56.

Serra, T., Zilberman, D., Goodwin, B.K and Featherstone, A. (2006). Effects of decoupling on the mean and variability of output. *European Review of Agricultural Economics* 33: 269-288.

Serra, T., Zilberman, D., Goodwin, B.K and 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.

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

Wansbeek, T. and Kapteyn, A. (1989). Estimation of the Error-Components Model with Incomplete Panels. *Journal of Econometrics* 41: 341-361.

Wooldridge J. M. (2001). *Econometric Analysis of Cross Section and Panel Data*. Cambridge: The MIT Press.