Extended Abstract

Please do not add your name or affiliation

Paper/Poster Title	***Explanation of Persistent and Transient efficiency in the presence of endogeneity – a one-step approach without distributional assumption: the case of Hungarian crop producing farms***

Abstract prepared for presentation at the 98th Annual Conference of The Agricultural Economics Society will be held at The University of Edinburgh, UK, 18th - 20th March 2024.

Abstract	***200 words max***

In this paper, we combine advances from two recent paper to explain persistent (PTE) and transient (TTE) efficiency in the presence of endogeneity. Earlier methods applied either a multi-step or a one-step procedure to estimate PTE and TTE. Until now, it was possible only with multi-step procedure to account for endogeneity in the context of PTE and TTE estimation. The virtue of our proposed method that it allows the estimation of the determinants of PTE and TTE in the presence of different sources of endogeneity applying a one-step approach without distributional assumption on the one-sided error term.

The paper also contributes to the empirical literature examining the impact of crop diversity and size of farms on TTE and PTE using Farm Accountancy Data Network (FADN) data for Hungarian crop producers for the period (2013 - 2021).

Preliminary results show that the proposed method give reasonable estimates of the production technology and persistent and transient efficiency determinants. In addition, the results suggest that in order to get consistent estimates it is important to account for different sources of endogeneity in the case of Hungarian crop producing farms.

Keywords	persistent technical efficiency (PTE) transient technical efficiency (TTE), four component model; endogeneity; Hungarian agriculture
JEL Code	Q12 see: www.aeaweb.org/jel/guide/jel.php?class=Q)

Introduction	***100 – 250 words***

Recent developments in the stochastic frontier (SF) literature have focused on coping with the problem of biased estimates of technology parameters and technical efficiency arising from various sources of endogeneity and with issue of heterogeneity. Colombi et al. (2014) and Kumbhakar et al. (2014) have developed a framework for estimating a four-component panel SF model, which model separates firm-specific time-invariant heterogeneity from persistent and transient technical inefficiency and the stochastic noise term.

In a recent paper (Danelon et al., 2021) proposed a deterministic frontier approach to estimate the four component frontier model. The main novelty of their proposed approach that no distributional assumptions are needed for the estimation of PTE and TTE. As the authors of this paper apply a cost function, they argue that the firm effect is unlikely to be correlated with the argument of the cost function (input prices and outputs) which are exogenous to the firms,

therefore they did not account for endogeneity in their model. However, this later assumption is unlikely for production functions.
Bokusheva et al., 2022 proposed a multi-step procedure in the case of the four-component model that accounts for three sources of potential endogeneity: (i) unobserved heterogeneity; (ii) simultaneity of input use with both types of technical efficiency; (iii) potential correlation of the noise term with the regressors.
Our aim is to combine the above two methods. As a result, we will be able to estimate the four-component model applying a one-step approach and at the same time to account for the above mentioned three sources of potential endogeneity.

Methodology	***100 – 250 words***

Our starting point is the four-component SF model developed by Colombi et al. (2014) and Kumbhakar et al. (2014). The model in general might be written as follows:

$$y_{it} = \alpha_0 + \boldsymbol{\beta}'\boldsymbol{x}_{it} + \chi_i + v_{it} - \eta_i - u_{it} \quad (1)$$

where v_{it} and u_{it} are i.i.d. variables following $N(0, \sigma_v^2)$ and $N^+(0, \sigma_u^2)$, respectively; χ_i is assumed to be i.i.d. $N(0, \sigma_\chi^2)$ and η_i is i.i.d. $N^+(0, \sigma_\eta^2)$. Further, $\eta_i + u_{it}$ is defined as overall technical inefficiency. In this model formulation, all components of the error term are assumed to be independently distributed of each other and of the regressors.
Two different estimation procedure has been proposed in the literature to estimate the model in 1. The One-step approach that use very complex log-likelihood function formulations (Colombi et al. 2014; Lai and Kumbhakar 2018) and the multi-step approach that was proposed by Kumbhakar et al. (2014).
In this paper following Danelon et al. (2021), we do concentrate on the explanation of PTE and TTE instead of estimating efficiency levels. In this case we do not need to make any distributional assumptions on η_i and u_{it}. Instead of this we can assume that $E(\eta_i) = h_1(z_i)$ and $E(u_{it}) = h_2(w_{it})$, where z_i and w_{it} are potential determinants of PTE and TTE. With these assumptions the model can be estimated in one step applying a non-linear generalised method of moments (GMM) estimator.

Results	***100 – 250 words***

Results show that the proposed method give reasonable estimates of the production technology and persistent and transient efficiency determinants. In addition, the results suggest that in order to get consistent estimates it is important to account for different sources of endogeneity in the case of Hungarian crop producing farms.
Results also suggest that crop diversity and size of farms have different effect on TTE and PTE. Size have a positive effect on PTE, while it did not have any significant effect on PTE. Crop diversity affected positively TTE, but its effect was not significant on PTE.

Discussion and Conclusion	***100 – 250 words***

The main novelty of the paper that it propose an approach that allows the estimation of the four-component model in one-step and account for three sources of potential endogeneity.
One additional advantage is that no distributional assumptions are needed for estimation of the persistent and transient efficiency components. Instead of the assumptions, that the persistent and transient efficiency are random, they are specified as deterministic functions. However, a relative disadvantage of this approach that it is not possible to identify the constant terms in the persistent and transient efficiency components and therefore, only the estimation of relative

efficiencies are possible. However, for purposes of empirical estimations and policy implications it is not a significant disadvantage. Policy makers and stakeholders are usually more interested in the determinants of (in)efficiency than in the absolute values of efficiency levels. The determinants of in(efficiency) can be easily obtained within this approach, because the marginal effects of Z_i^{η} and Z_{it}^{u} on (in)efficiecy can be estimated without knowing their absolute level and the marginal effects of Z_i^{η} and Z_{it}^{u} on the absolute and relative efficiency are the same.