https://doi.org/10.17221/2/2020-AGRICECON

The effects of the participation in producer organisations on the performance of dairy farmers in the Czech Republic and future challenges

*Iveta Boskova[1], Samuel Ahado[2], Tomas Ratinger[2,3]**

[1]Institute of Agricultural Economics and Information (IAEI), Prague, Czech Republic
[2]Department of Economics and Development, Faculty of Tropical AgriScienecs, Czech University of Life Sciences (CULS), Prague, Czech Republic
[3]Technology Centre of the Czech Academy of Sciences (TC CAS), Prague, Czech Republic

**Corresponding author: ahado@ftz.czu.cz*

The authors are fully responsible for both the content and the formal aspects of the electronic supplementary material. No editorial adjustments were made.

Electronic supplementary material

Supplementary Material S1

TFP index definition

We used Törquist-Theil index of total factor productivity (TFP) of a farm relatively to the average farm in the base year. Törquist-Theil TFP is based on the theory of index numbers; it is a useful alternative for econometric estimates of efficiency when data do not allow to estimate the respective production function (Capalbo and Antle 1998). TFP index is defined as a ratio of the production and input indexes (relative change of the production of outputs over relative change of the use of inputs).

Production and input indices are calculated as weighted averages of changes of produced outputs or used inputs, where weights are given by shares of outputs or inputs (factors) in total revenue or total cost respectively.

In the case of two farms (firms) i and b, which produce n outputs Q_j ($j = 1, \ldots, n$) by using m inputs X_k ($k = 1, \ldots, m$). The index $t = \ln TFP$ is defined as follows:

$$t_1 = \frac{1}{2}\sum_{j=1}^{n}\left(R_j^i + R_j^b\right)\left(\ln Q_j^i - \ln Q_j^b\right) - \frac{1}{2}\sum_{k=1}^{m}\left(S_k^i + S_k^b\right)\left(\ln X_k^i - \ln X_k^b\right) \quad (1)$$

where: for a farm (firm) i, R_j^i – the share of the revenue j on the total revenue of all n outputs and S_k^i – the share of the k^{th} input on the total cost of all m inputs.

Direct Covariate Matching

To understand the effect of treatment (in our case participating in a cooperative) we have to choose or to construct a control farm with "identical" characteristics from the pool of non-participating producers. The standard framework for counterfactual analysis provides Roy-Rubin-model (Caliendo and Kopeinig 2008). In this model, we recognise three types of treatment effects: average treatment effect on treated (*att*), average treatment effect on controls (*atc*) and average treatment effect (*ate*) which averages the both former. *ate* is defined:

https://doi.org/10.17221/2/2020-AGRICECON

$$\tau_{ate} = E\left[Y(1) - Y(0)\right] \tag{2}$$

where: $\tau = Y(1) - Y(0)$, $Y(D)$ is an result variable; D equals 1 if the unit is a member of a cooperative (treatment) and 0 otherwise.

The sample *ate* (*sate*) takes the form of:

$$\tau_{sate} = \frac{1}{N}\sum_{i}\left[Y_i(1) - Y_i(0)\right] \tag{3}$$

Matching estimators are based on imputing a value on the counterfactual outcome for each unit. Abadie and Imbens (2002), propose direct matching which is based on Mahalanobis metric $||x|| = (x'Vx)^{1/2}$; where x – vector of structural variables and V – a positive semidefinite matrix. This metric is used to determine the nearest similar unit(s). Let M denotes the number of nearest matched unit to unit i (in the paper we used $M = 4$) and $K_M(i)$ denote the number of times i is used as a match for all observations of the opposite treatment group, each time weighted by the total number of matches for observation. Abadie and Imbens (2002) showed that the sample average treatment effects (*satt*, *satc*, *sate*; *s* denotes "sample") and their variance depend on $K_M(i)$. For example the simple estimator of *ate* takes the following form:

$$\tau_{sate} = \frac{1}{N_1}\sum_{i=1}^{N}(2D-1)\left[1 + K_M(i)\right]Y_i \tag{4}$$

Such simple estimators will be biased in the finite set if the matching is not exact. Abadie and Imbens (2002) therefore propose using regression estimates of Y as a linear function of the considered structural variables (covariates). for *satt* in the control group $\hat{\mu}_0(x)$, for *satc* in the sub-sample of participants $\hat{\mu}_1(x)$ and for *sate* using the both regressions. The bias adjusted *sate* given by the regression estimates looks as follows:

$$\tilde{\tau}_{sate} = \frac{1}{N}\sum_{i}\left(\tilde{Y}_i(1) - \tilde{Y}_i(0)\right) \tag{5}$$

It is also important that Abadie and Imbens (2002) derived a robust estimator of variance which is essential for the statistical judgement of the effects. For more information, look in Abadie and Imbens (2002) or Ratinger et al. (2013).