Analysis of Production Function of Agricultural Holdings

Marie Šimpachová Pechrová[1], Ondřej Šimpach[2]

Abstract. Agricultural holdings differ in their production amount due to the size, different yields, technology or weather conditions. The aim of the paper is to find appropriate model for modelling the production function that would account for this heterogeneity and to find the main determinants of the agricultural production. There are panel data available (2 268 observations for 2014–2017), so fixed versus random effects model are considered, compared and tested to choose appropriate model. Cobb-Douglas production function was modelled. The production of farm (sales of own products and services) was explained by the consumption of material, capital, number of employees, and acreage.

Results of fixed effects model were not economically meaningful. Hausmann test proved that random effects model is preferred. Increase of consumed material by 1% caused the increase of production by 0.52%, of tangible long-term assets by 1% caused increase in production by 0.17%, increase of number of employees by 1% caused increase in production by 0.01%, and of land by 0.02%. The production of Czech agricultural holdings is influenced the most by the amount of consumed material, while the number of employees and acreage are less significant.

Keywords: agriculture, fixed effects model, panel data, random effects model

JEL Classification: C23, C51, C12
AMS Classification: 62J05, 62F03

1 Introduction

Agricultural holdings differ in their production amount not only due to the size, but also because of different yields caused by technology or weather conditions. The level of used technology and innovations on the farm is a prerequisite of its level of production and competitiveness. Technology transfer between universities and public or private research institutes and producers can be hence seen as a driver of innovative entrepreneurship in agriculture and the agri-food industry. (Carayannis, Rozakis and Grigoroudis [4]).

The weather changes are particularly important in crop production, but also livestock production can experience the fluctuation in production, especially when pasture system is concerned. For example, Perez-Mendez, Roibas and Wall [16] showed that meteorological variables had a significant impact on milk production (it is higher under warm conditions due to improvements in forage production) in Spanish region of Asturias. Observed and unobserved variables influence the production of agricultural holdings. Therefore, mathematic modelling is useful tool that can help decision making and planning of farmers.

For example, in milk production in New Zealand, a multi-stage stochastic program that incorporates weather dynamics was introduced by Dowson et al. [7] in order to help the decision making and planning of the farmers. Duc-Anh et al. [8] used seasonal climate forecasts in the sugar industry in Australia and demonstrated, that the model can facilitate the irrigation planning of sugarcane farming and consequently contribute to the improvement of gross margin of the growers. Nechaeva et al. [14] identified on a case study of Russia's region Nizdhny Novgorod the factors that influence effective operation of agricultural organizations in terms of their decision making about equipment and machinery. "The research identified two major factors providing a competitive market position of manufacturing companies: factor 1 *Technical and technological modernization of production* and factor 2 *Qualitative composition safety*" (Nachaeva et al., 2019). Daxini et al. [6] examined how the psychological factors can influence the decision making of farmers about adoption of nutrient management practices in Ireland by the behavioural model. The results showed that (on a national sample) attitudes, subjective norms (social pressure), perceived behavioural control (ease / difficulty) and perceived resources were significant and positively associated with farmers' intentions.

[1] Institute of Agricultural Economics and Information, Department of Modelling of the Impacts of Agricultural Policy, Mánesova 75, 120 00 Prague 2, simpachova.marie@uzei.cz.
[2] Institute of Agricultural Economics and Information, Independent Department of Operational Analysis and International Cooperation, Mánesova 75, 120 00 Prague 2, simpach.ondrej@uzei.cz.

Certain effect on production can have also agricultural subsidies provided by the European Union, despite that they shall be decoupled from the amount of production (are not provided on unit of production, but rather on hectare of agricultural land). Pechrová [15] examined to what extend EU's agricultural subsidies influence the production based on the production function of the organic farms in the Czech Republic. She found out that according to the expectation the increase of material, capital, labour and land cause the increase of production and Single Area Payment Scheme (SAPS) and other subsidies mildly decrease the production. Hence, their effect is not significant as it shall be.

The aim of the paper is to find appropriate model for modelling the production function of agricultural holdings that would account for this heterogeneity and to find the main determinants of the agricultural production. There are panel data available, so fixed versus random effects model are considered, compared and tested in order to choose appropriate model. The selection between random and fixed effects model was done for example by Hu and McAleer [11] on a panel data of agricultural production in China. Dataset consisted of 30 provinces for the period 1991–1997. They estimated Cobb-Douglas production function using both, fixed and random effects models, and calculated provincial technical inefficiency. Testing revealed that for this application the random effects model was rejected in favour of the fixed effects model. Also, Heshmati [9] compared the two above mentioned models using a sample of rotating panel data set from Swedish crop producers over the period 1976–1988. He showed that the introduction of heteroscedasticity and the integration of sample selection in the production relationship was important. "The impact of a correction for selectivity bias on the results, in terms of input elasticities and returns to scale is found to be significant" (Heshmati [9]).

2 Data and methods

The aim of the paper is to find appropriate model for modelling the production function of agricultural holdings on panel data. The advantage of panel data is that they can control for unobservable heterogeneity among farms (Ye, Xu and Wu [18]). They allow to control for unobservable variables or variables that change over time but not across entities. "A major motivation for using panel data has been the ability to control for possibly correlated, time-invariant heterogeneity without observing it." (Arellano [1]). "Panel data give more informative data, more variability, less collinearity among the variables, more degrees of freedom and more efficiency. Panel data are better able to study the dynamics of adjustment and are better able to identify and measure effects that are simply not detectable in pure cross-section or pure time-series data." (Baltagi [2]) Some drawbacks might be difficulties with data collection.

Our analysis utilises accounting data from Albertina database collected by Bisnode company from public firms' register. Acreage of land was obtained from Land Parcel Identification System. There were 2 268 observations for 680 companies for years 2014 to 2017. There were on average 3.3 observations per one firm. All calculations were done in econometric SW Stata version 15.1.

A Cobb Douglas type of production function was constructed where explained variable – the production of farm *i* in thousands of CZK (y_{it}) – was represented by the sales of own products and services in particular year (*t*). In order to remove the impact of price changes overtime, the production was deflated by the price index of agricultural producers (2010 = 100).

Explanatory variables included material, capital, labour, and land. Material ($x_{1,it}$) consisted of consumed material and energy by i^{th} farm in time *t* and capital ($x_{2,it}$) of tangible long-term assets of i^{th} farm in time *t*. Both variables were deflated by the industrial producers' price index (2010 = 100). Labour ($x_{3,it}$) was represented by number of employees. The acreage of farmland (input land – $x_{4,it}$) was obtained from LPIS database.

A Cobb-Douglas form of production function was selected (1)

$$y_{it} = \beta_0 x_{1,it}^{\beta_1} x_{2,it}^{\beta_2} x_{3,it}^{\beta_3} x_{4,it}^{\beta_4} \tag{1}$$

where β_0 is constant, $\beta_{1,\,2,\,\ldots 4}$ are coefficients of explanatory variables and estimated in linearized form (2):

$$\ln y_{it} = \ln \beta_0 + \beta_1 \ln x_{1,it} + \beta_2 \ln x_{2,it} + \beta_3 \ln x_{3,it} + \beta_4 \ln x_{4,it} + u_{it} \tag{2}$$

If the panel nature of the data is ignored, all data can be pooled together, and the model can be estimated by simple ordinary least square method (OLS). Therefore, firstly a Breusch-Pagan Lagrange multiplier test was used to decide between a random effects regression and a simple OLS regression. The null hypothesis (H_0) is that variances across farms are zero, there is no significant difference across farms (i.e. no panel effect). If the null hypothesis is rejected, there is random effect present in a model. Both, fixed and random effects models were estimated and compared based on Hausmann test.

Fixed effect model (FEM) requires that all variables change over time otherwise it removes the effect of time-invariant characteristics. Each farm has its own intercept because each farm has its own individual characteristics that may influence the explained variable such as the management, sowing method etc., while the slope coefficient is the same for all. It is assumed that some feature of individual farm may impact or bias the predictor or outcome variables. It is assumed that there is a correlation between farm's error term and predictor variables, but time-invariant characteristics unique to the farm are uncorrelated among themselves. If the error terms are correlated, then FEM is no suitable since inferences may not be correct.

In **random effects model** (REM) the variation across entities is assumed to be random and uncorrelated with the predictor or independent variables in random effects model. It is assumed that differences across farms have an influence on dependent variable. This allows for time-invariant variables to play a role as explanatory variables.

Consequently, **Hausman test** was done to decide between fixed or random effects. It tests whether the unique errors are correlated with the regressors. The null hypothesis is that they are not (difference in coefficients is not systematic), which means that preferred model is random effects vs. the alternative fixed effects. If value W of the Hausman test which follows χ^2 distribution is significant (p-value < 0.05), then random effects model shall not be used. If the p-value is higher than $\lambda = 0.05$, H_0 is rejected and random effect model shall be used.

The **goodness of fit** was assessed by coefficient of multinomial determination. It is calculated as $R^2 = 1 - (\sigma_{ui}^2/\sigma_{yi}^2)$, where σ_{ui}^2 is residual variance and σ_{yi}^2 is total variance in a model.[3]

Statistical **significance of the parameters** was tested by t-test with H_0: the coefficient is not statistically significantly different from 0. Test criterion is calculated as the ratio of the parameter's absolute value (β_k) to the standard error (S_{bi}). If the p-value is lower than $\lambda = 0.05$, H_0 is rejected.

Coefficient of intraclass correlation (ρ) tells that how much of the total variance is given by the differences across panel. It is calculated as the ratio $\sigma_{ui}^2/(\sigma_{ui}^2 + \sigma_{ei}^2)$, where σ_{ui}^2 is variance of residuals within groups and σ_{ei}^2 variance of overall error term.

3 Results

First, FEM was estimated using Stata command `xtreg` and `areg`. Results are displayed in Table 1. Variable land was omitted due to collinearity (it does not change overtime). We considered fixed-effects (within) regression model where there are n entity (firm) – specific intercepts. Correlation between error u_i and the regressors ($\mathbf{X}_b$) was assumed and confirmed in the results (regression coefficient is 0.74). The model is statistically significant as p-value of F test ($F^{[3,\ 1585]} = 102.15$) indicates that the null hypothesis (H_0: all coefficients in the model are jointly equal to zero) is rejected. Also p-value of F-test for testing that all within groups residues u_i are equal to 0 (value $F^{[679,\ 1585]} = 23.93$) showed that H_0 is rejected. At least one error is different from zero. According to ρ 95.40% of variance is due to differences across panels.

The value of the coefficients does not correspond to the expectations. Two-tail p-values of t-test for each coefficient shows that only constant and consumption of material and energy are statistically significant. If it increases by 1%, then the production increases by 0.34%. Capital is insignificant determinant of the production as same as the number of employees. The second mentioned coefficient is even negative, indicating possible overemployment when the increase of the number of employees cause decrease of production. Hence, the model itself is not optimal.

Another way of estimation of fixed effect model was done to compute real R^2 value. It has again n entity-specific intercepts and the results are similar. R^2 was 98.22% and after adjustment 97.46% which indicates high goodness of fit of the model. This means that the variability of explained variable is explained from 97.46% by variability of explanatory variables.

Because the results of fixed effects model were not satisfying, random effects model was estimated too. First, it was tested whether to use random effects model or simple OLS regression by Breusch and Pagan Lagrangian multiplier test. χ^2 value of the test was 1476.77 with p-value 0.00. Hence, we rejected H_0. There are significant differences among the farms and random effect model is appropriate. Simple OLS regression cannot be done.

[3] The command `xtreg` in SW Stata 15.1 provides meaningless coefficient of multinomial determination (R^2) and adjusted R^2. Therefore, command `areg` was utilized to report correct values of R^2.

Variable	Coef.	Std. error	t-value	p-value
Constant	6.8437	0.2715	25.2100	0.0000
ln x1	0.3415	0.0199	17.1600	0.0000
ln x2	0.0183	0.0225	0.8100	0.4160
ln x3	-0.0018	0.0012	-1.4900	0.1360
ln x4	(omitted)			
σ_u	1.3642		σ_e	0.2996
ρ	0.9540		corr (u_i, $\mathbf{X}_b$)	0.7359
$F^{[3, 1585]}$	102.1500		p-value (F)	0.0000
$F^{[679, 1585]}$	23.9300		p-value (F)	0.0000
R^2	0.9822		$R^2_{adjusted}$	0.9746

Table 1 Results of fixed effects model

Therefore, random effects model was estimated. Results are displayed in Table 2. The model as a whole was statistically significant as the probability of Wald χ^2 = 2703.09 was 0.00. All estimated coefficients were statistically significant too, but again with the exception of the number of employees. We can conclude that the number of employees does not statistically significantly influence the amount of production of agricultural companies in a sample as both models considered it statistically insignificant. Maybe due to seasonality of the work in agriculture, the labour shall be measured in Annual Work Units (AWU) to be statistically significant determinant but our dataset does not contain it. Besides, agriculture is labour demanding mainly in case of livestock production. Crop production requires less employees and (mostly) seasonal. The sample is mixed but might be with the predominance of crop production. Values of other coefficients corresponds to the expectations and are statistically significant. Increase of consumption of material and energy by 1% brings increase of production by 0.52%, the elasticity is the highest here. Then the increase of capital by 1% brings increase of production by 0.17%. Increase of acreage of land by 1% brings increase of production by 0.02%. The lowest elasticity is in case of number of employees, but it is not a significant determinant.

Variable	Coef.	Std. error	t-value	p-value
Constant	3.5058	0.1495	23.4600	0.0000
ln x1	0.5217	0.0150	34.8500	0.0000
ln x2	0.1704	0.0167	10.2000	0.0000
ln x3	0.0008	0.0013	0.6600	0.5110
ln x4	0.0193	0.0057	3.4100	0.0010
σ_u	0.8693		σ_e	0.2996
ρ	0.8938		corr (u_i, $\mathbf{X}_b$)	0 (assumed)
Wald $\chi^{2[4]}$	2703.09		p-value (F)	0.0000
R^2	0.7948			

Table 2 Results of random effects model

Hausman tests ($\chi^2_{[3]} = (b-B)^T \left[V_b - V_B\right]^{-1} (b-B)$ = 216.87, p-value = 0.00) showed that null hypothesis is rejected, the difference in coefficients are not systematic, and fixed effects can be used. Farms' errors (u_i) are not correlated with the regressors, which is the assumption of fixed effects model. Hence, random effect model is more appropriate to model the relation between production and production factors. This model also complies with economic assumptions set prior the estimation. Also, random effects model can deal with regressors that are fixed across individual farms (land in our case). Random effects models offer distinct advantages over fixed effects models in terms of their efficiency and their ability to calculate shrunken residuals (Clarke et al. [5]).

Also, the total residual variance can be partitioned into two components: the between-farm variance σ_u^2 and the within-farm variance σ_e^2. In our case, the variance between farms in higher (0.87) than within farm (0.30). According to ρ 89.38% of variance is due to differences across panels.

On the other hand, there are certain drawbacks. There is a certain probability that there is a correlation between the unobserved effects and the explanatory variables. In fixed effects model it is assumed that there is certain correlation, but in random effects model is considered to be zero, that does not have to be true. This implies inconsistency due to omitted variables. On the other hand, fixed effects model is consistent despite being inefficient.

The correlation problem can be overcome by fitting correlated random-effects and correlated random-coefficient models. It is used for example by Suri [17] on the decision making of Kenyan farmers to adopt new technologies where the benefit and costs of technologies are heterogenous, so the farmers with low net returns do not adopt new technology. It his case, the returns to new technology (hybrid wheat) are heterogeneous and correlated with the decision to use hybrid, so he estimates correlated random coefficient model. For an ease of the calculation, there is a community-contributed command `randcoef` in Stata programmed and explained by Cabanillas, Michler and Michuda et al. [3].

Unobserved heterogeneity in panel data modelling concerned Hsiao [10]. He used various formulations of the interactive effects models and suggested a quasi-likelihood approach as a framework to study issues of estimation and statistical inference when regressors are either strictly exogenous or predetermined and under different combinations of the data size. Also, Musafiri and Mirzabaev [13] noted that the results of their cross-section model shall be treated with caution due to unobserved heterogeneity among agricultural households in Rwanda which they surveyed. Therefore, they constructed also panel data model – particularly pooled OLS (difference in difference) model and fixed effects model to observe the changes in agricultural output over time. "The results confirm the predominant role of labour, capital, and land quality to output growth over time." (Musafiri and Mirzabaev [13]).

It is therefore reasonable to assume that production factors and other differences across farms have an influence on production because each farm has slightly different structure of production, is located in different weather conditions and is managed variously. All those unobserved characteristics can affect the volume of production. For example, it was proved by Perez-Mendez, Roibas and Wall [16] that meteorological variables can affect milk production through their impact on the productivity of cows and the production of foodstuff. They proposed a production function where meteorological variables affect the productivity of cows and the production of forage, thereby indirectly affecting milk production.

Mundlak and Butzer [12] analysed the role of capital in agricultural production in the cross-country study of production functions. The results of their model revealed the relative importance of capital, because the lack of it can be a constraint on agricultural growth. In our case, the highest intensity on production had the consumption of material and energy. The coefficient (which is also an elasticity coefficient) shows that 1% change of the consumption causes 0.52% change in production. Capital and land changes are less important. The lowest elasticity had the labour which was even statistically insignificant. It can be due to the fact that "The shift to more productive techniques is associated with a decline in labor, reflecting labor-saving technical changes" (Mundlak and Butzer [12]).

Conclusion

Agricultural holdings differ in their production amount not only due to the size, but also because of different type of production, yields caused by technology or weather conditions. The aim of the paper was to find appropriate model for modelling the production function of agricultural holdings that would account for this heterogeneity and to find the main determinants of the agricultural production. Fixed versus random effects model were considered, compared and tested in order to choose appropriate model. A sample of agricultural holdings consisted of 2 268 observations for 680 groups of farms for years 2014–2017. Production function was modelled in Cobb-Douglas form that has an advantage that the coefficients can be interpreted as elasticities. Production was explained by material, capital, labour, and land. Accounting data were taken from Albertina database. The acreage of farmland was obtained from LPIS land register.

Then we tried to decide whether to use fixed effects or random effects model to estimate the production function of agricultural firms. The nature of the data would suggest using random effects model. It is due to the fact, that the acreage of land does not change during the observed period of time. For this reason, this variable will be omitted from fixed effects model. However, fixed effects model still might be better to estimate the production function of agricultural holdings.

Results provided by fixed effects model were not economically feasible. Coefficient in case of labour was negative suggesting that increase of number of employees would bring decrease in production. Probably AWU would be more suitable variable, but it was not available for us. Also, as the acreage has not changed much during years, it was omitted from the calculation.

Hausmann test also proved that random effects model is preferred to fixed effects. Therefore, a random effects model was estimated. The results corresponded to the economic theory that increase of consumed material by 1% cause the increase of production by 0.52%, increase of usage of tangible long-term assets by 1% cause the increase of production by 0.17%, increase of number of employees by 1% increase the production by 0.01% and increase of land by 1% cause increase by 0.02% in production. The production of Czech agricultural holdings is influenced the most by the amount of consumed material, while the number of employees and acreage play minor role (there were 11% of farms with only livestock production).

Acknowledgements

The paper was supported by the Ministry of Agriculture of the Czech Republic, institutional support MZE-RO0918, Internal Research Project no. 1113/2019.

References

[1] Arellano, M. (2003). *Panel Data Econometrics*. New York: Oxford University Press. ISBN 0-19-924528-2.
[2] Baltagi, B. H. (2005). *Econometric analysis of Panel Data.* New York: John Wiley & Sons Inc. ISBN 0-470-01456-3.
[3] Cabanillas, O. B., Michler, J. D. & Michuda, A. et al. (2018). Fitting and interpreting correlated random-coefficient models using Stata. *Stata Journal*, ***18***(1), 159–173.
[4] Carayannis, E. G., Rozakis, S. & Grigoroudis, E. (2018). Agri-science to agri-business: the technology transfer dimension. *Journal of Technology Transfer*, ***43***(4), 837–843. DOI: 10.1007/s10961-016-9527-y.
[5] Clarke, P., Crawford, C., Steele, F. & Vignole, A. (2010). The Choice Between Fixed and Random Effects Models: Some Considerations for Educational Research. *Discussion Paper No. 5287*. IZA DP No. 5287. Available from http://ftp.iza.org/dp5287.pdf.
[6] Daxini, A., O´Donoghue, C., Ryan, M., Buckley, C., Barnes, A. P. & Daly, K. (2018). Which factors influence farmers' intentions to adopt nutrient management planning? *Journal of Environmental Management, **10***, 350–360.
[7] Dowson, O., Philpott, A., Mason, A. & Downward, A. (2019). A multi-stage stochastic optimization model of a pastoral dairy farm. *European Journal of Operational Research*, ***274***(3), pp. 1077–1089. DOI: 10.1016/j.ejor.2018.10.033.
[8] Duc-Anh, A., Mushtaq, S., Reardon-Smith, K., Kouadio, L., Attard, S. & Cobon, D. (2019). Value of seasonal forecasting for sugarcane farm irrigation planning, *European Journal of Agronomy*, ***104***, 37–48. DOI: 10.1016/j.eja.2019.01.005.
[9] Heshmati, A. (1994). Estimating random effects production function models with selectivity bias: an application to Swedish crop producers. *Agricultural Economics, **11***(2–3), 171–189. DOI: 10.1016/0169-5150(50)00017-3.
[10] Hsiao, Ch. (2018) Panel models with interactive effects. *Journal of Econometrics*, ***206***(2), 645–673.
[11] Hu, B. & McAleer, M. (2005). Estimation of Chinese Agricultural Production Efficiencies with Panel Data. *Mathematics and Computers in Simulation*, ***68***(5–6), 474–483.
[12] Mundlak, Y. & Butzer, R. (1997). The Determinants of Agricultural Production: A Cross-Country Analysis. *Policy Research Working papers.* The World Bank. https://doi.org/10.1596/1813-9450-1827
[13] Musafiri, I. & Mirzabaev, A. (2014). Long-Term Determinants of Agricultural Output in Smallholder Farmers in Rwanda. In *Proceedings of Agricultural & Applied Economics Association's 2014 AAEA Annual Meeting* (pp. 1–39). Minneapolis: Agricultural & Applied Economics Association.
[14] Nechaeva, M. L., Nechaev, V. N., Ilicheva, O. V., Kozlova, L. A. & Kurilova, A. A. (2019). Material and Technical Base of Agricultural Organizations as the Basis for their Economic Efficiency. *International Journal of Advanced Biotechnology and Research*, ***10***(1), 363–370.
[15] Pechrová, M. (2014). The Impacts of the EU's Subsidies on the Production of Organic Farms. In Jedlička, P. (Ed.), *Proceedings of Hradec Economic Days.* Hradec Králové: University of Hradec Králové, 124–130. ISSN 2464-6067.
[16] Perez-Mendez, J. A., Roibas, D. & Wall, A. (2019). The influence of weather conditions on dairy production. *Agricultural Economics*, ***50***(2), 165–175. DOI: 10.1111/agec.12474.
[17] Suri, T. (2011). Selection and Comparative Advantage in Technology Adoption. *Econometrica*, ***79***(1), 159–209.
[18] Ye, X., Xu, J. & Wu, X. (2018). Estimation of an unbalanced panel data Tobit model with interactive effects. *Journal of choice modelling*, ***28***(C), 108–123.