

Heterogeneous Technology, Scale of Land Use and Technical Efficiency: the Case of Hungarian Crop Farms

Lajos Baráth^a, Imre Fertő^{ab}

^aInstitute of Economics, Centre for Economic and Regional Studies, Hungarian Academy of Sciences, Budaörsi út 45, 1112 Budapest, Hungary ^bCorvinus University of Budapest, Fővám tér 8, 1093 Budapest, Hungary

Abstract

In this paper, we analyse the technical efficiency (TE) of Hungarian crop farms between 2001 and 2009 using conventional stochastic frontier analysis (SFA) and a latent class model (LCM). Our results suggest that technological heterogeneity plays an important role in Hungarian cereal, oilseed and protein (COP) crop producing farms, which are traditionally assumed to use homogeneous technologies. We identify two groups with different technologies using an LCM, and find that the scale of land use is one of the most important factors that differentiates our sample into unique technologies. Our results reveal that in the Hungarian COP sector there is less chance to increase performance through TE improvement than was earlier expected. The estimations also suggest that there is no room to improve productivity by increasing farm size, unless farms switch technologies. Consequently, agricultural policies for increasing productivity should concentrate on technological progress.

Keywords: latent class model; technical efficiency; crop farm; technological heterogeneity; scale of land use

1. Introduction

A growing body of literature has emphasised the importance of different technologies in estimating technical efficiency (TE) and technological change (TCH) in the dairy sector (Alvarez-Del Corral, 2010; Alvarez et al., 2012; Emvalomatis, 2007; Sauer-Morrison, 2013). These studies generally agree that the traditional frontier models, which do not account for heterogeneous technologies, produce biased TE and TCH estimates and suggest that latent class models (LCMs) should be used to control for technological heterogeneity.

Although LCMs have produced promising results in the dairy sector, they have not been used to examine the crop sector. This raises the question of whether LCMs are adequate for the crop sector. The main goal of our paper is to investigate this question. Therefore, we aim to contribute to the existing literature mainly from a methodological point of view.

Nevertheless, our results have important policy implications, for the following reasons. Increasing productivity has been one of the central aims of the common agricultural policy (CAP) since its foundation. Although policy makers can use different policy measures to improve productivity, selecting an effective policy measure requires consistent estimates of the components of productivity change (O'Donnell, 2012). However, biased estimates can lead to inadequate policy implication. For instance, if TE is higher, as papers that use models which are able to handle the heterogeneity bias generally suggest (Alvarez-Del Corral, 2010, Alvarez et al, 2012, Emvalomatis, 2007, Sauer-Morrison, 2013), then policies that are designed to improve productivity through TE (e.g. through education or extension programs) may be less effective thereby requiring decision makers to choose different measures. One possible option is to increase productivity through technological progress (e.g. through investment in R&D). However, if the technology is heterogeneous, the traditional models also yield biased TCH measures, as Sauer-Morrison (2013) suggests. In particular, the presence of different technologies in an industry makes the empirical analysis of technological change more complex than is typically modelled by shifts in a common production frontier. Therefore, in this paper we take into account the effects of heterogeneous technologies, and our results may help policy makers to choose suitable policy measures.

Technological differences and their effects on TE could be especially important in the Central and Eastern European transition countries (CEECs). Following regime change, a dual agricultural structure evolved in the CEECs (Bignebat and Latruffe, 2010; Ciaian et al., 2009; Csáki and Lerman, 1999), with a large number of small-scale farms on one side and a small number of large-scale farms on the other. It is assumed that the farms in this polarised environment use different technologies. Thus, we focus on Hungarian COP crop producing farms, which have a typical dual structure, in our empirical analysis.

The remainder of this paper is organised as follows. We briefly review the main changes that have occurred in Hungarian agriculture in the last decade, and then examine the efficiency of CEECs agriculture. Next, we outline the methods used in the analysis and present our dataset and then present our empirical results on TE. Finally, we conclude the paper and derive a number of policy implications.

2. Main characteristics and development of Hungarian specialised fieldcrop producing farms in the last decade

Fieldcrop production has traditionally been a key sector in Hungarian agriculture. About 40% of all Hungarian farms specialise in fieldcrop production, and use 60% of the arable land and account for more than a third of the output of agricultural production (Pesti-Keszthelyi, 2010). This is the subsector of Hungarian agriculture that integrates well with international commerce, in that the product channels are well organised and the products comprise the largest proportion of agricultural exports (Pesti-Keszthelyi, 2010). Researchers in this field are mostly of the opinion that Hungary is competitive in the area of cereal production, but not in animal farming (Jámbor, 2009). The Hungarian fieldcrop sector underwent significant structural changes between 2000 and 2010, with the number of farms declining considerably (29%) (CSO, 2012). As background to this fall, the number of corporate farms has increased by 26%, whereas the number of individual farms has declined by 29%. Consequently, the average size of corporate farms has decreased by 30% to 352 hectares, and the mean size of individual farms has doubled to 6.2 hectares. These facts indicate that the Hungarian fieldcrop sector is still characterised by a dual structure.

Figure 1 illustrates the changes in the production volume of fieldcrop farms.

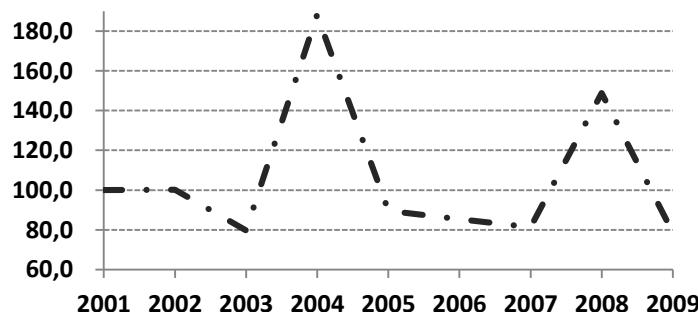


Figure 1: Development of cereal and industrial crop production output, 2001 = 100%

Source: Authors' illustration based on Hungarian Central Statistical Office (HCSO)

Figure 1 clearly shows that fieldcrop production volumes underwent considerable fluctuations during the period of analysis. No obvious tendency can be observed. Output volumes in the outstanding years were significantly higher than the output values in the average years; in 2004 alone, the volume of COP crop production was 80% higher than at the beginning of the 2000s. Several researchers have drawn attention to the effect that these substantial fluctuations had on farms' income (e.g., Harangi-Rákós-Szabó, 2010; Kapronczai, 2010). Kapronczai notes, 'These steep variations point to insufficiencies in soil nutrient levels as well as lower technology standards of Hungarian agriculture compared to agricultures in

developed countries' (Kapronczai, 2010, p. 77). This latter conclusion reinforces the importance of research on the technological level of farms, or of farms' TE.

3. Previous studies on efficiency and productivity in CEECs

There have been numerous efficiency analyses of agriculture in CEECs. Gorton and Davidova (2004) and Bojnec and Latruffe (2013) provided a detailed overview of TE papers in CEECs. In this section, we first present a short overview based on these reviews, and then mention some additional papers that analyse the TE of Hungarian agriculture and others that estimate the TE of CEECs by focusing on the issue of technological heterogeneity.

Gorton and Davidova (2004) gathered results from farm efficiency studies in six CEECs. Their main research interest was whether there is a clear superiority of a specific organisational type and the relationship between size and farm efficiency. They observed no clear evidence for the greater efficiency of one organisational form over another and they also did not find clear relationship between size and efficiency.

Bojnec and Latruffe (2013) reviewed papers analysing the influence of farm size and agricultural subsidies on farms' performance in CEECs. They found that farm size was positively related to farm efficiency in Slovenia and in the Czech Republic, whereas the relationship was ambiguous in Poland. Moreover, the authors found only one study analysing the effect of subsidies on farm performance in CEECs, namely Banse et al. (1999) showed a negative correlation between the level of protection and the level of performance in Hungarian agriculture.

Bakucs et al. (2010) used an SFA to evaluate the TE of Hungarian farms before and after accession to the European Union (EU), and investigated the determinants of efficiency. The results showed that EU membership reversed the pre-accession process of a decrease in efficiency. However, access to higher post-accession subsidies contributed to the lower efficiency of Hungarian farmers. Bakucs et al. (2012) then investigated the TE of Hungarian dairy farms between 2001 and 2008, again using an SFA. Their main finding was that the mean performance of individual and family farms was weaker than that of corporate farm organisations, including companies, cooperatives and intermediate and non-family farms, irrespective of method, product group and country. A simple mean comparison estimation revealed significant differences in farm performance among farms in terms of legal form and farm organisation. However, a panel regression only partially confirmed these results.

Latruffe et al. (2004) analysed TE and its determinants for a panel of individual farms in Poland specialising in crops and livestock. They found two important determinants of TE: farmers' education and the market integration of the farm.

Hockmann and Pieniadz (2007) applied the approach used by Alvarez et al. (2003) to consider farms' heterogeneity while exploring their (in)efficiency in Polish agriculture. The virtue of this method that it assumes a fifth, in the dataset unobserved firm-specific production factor in addition to land, capital, labour and intermediate input. In principle, this factor captures the effect of environmental conditions and covers differences in factor qualities such as climate condition, soil fertility and human capital, including management skills. The unobserved component is treated as a random variable and the model is estimated as a Random Parameter Model (RPM). Their results revealed the existence of the unobserved factor. Furthermore, their results confirmed that standard SFA models overestimate TE and their empirical findings also revealed that scale inefficiencies were not a severe problem in Polish agriculture.

Cechura (2010) estimated the TE of Czech agriculture using different SFA models and found that the average level of TE was around 90% for agricultural companies in the Czech Republic. Only those model specifications that allowed for the capture of time-invariant firm heterogeneity provide consistent estimates of TE.

In summary, recent studies have identified some of the potential factors affecting TE in CEECs, including farm size, farm organisation and policy measures. However, most of the studies on TE, except those conducted by Cechura (2010) and Hockmann and Pieniadz (2007), have assumed that the technology is the same for all farms. No study has analysed the effect of unobserved technological differences using an LCM in the crop sector. If unobserved technological differences are not taken into account, then TE scores may be underestimated. Our aim here is to clarify this issue.

4. Theoretical Background and Empirical Model

The two most common approaches to estimating technical efficiency are data envelopment analysis (DEA) and stochastic frontier analysis (SFA). DEA is a more flexible approach in the sense that it does not require the assumption of a special functional form, transforming inputs into outputs; however, it reacts very sensitively to outliers and inconsistencies in the data. In

contrast, SFA requires assumptions about the form of the production function and the distribution of inefficiencies.

Considering the high fluctuation in Hungarian agricultural output (see section 2), we suppose that the effect of stochastic shocks is significant, and thus use SFA. The stochastic production frontier was originally proposed by Aigner et al. (1977) and Meeusen and van den Broeck (1977), independently of each other.

A stochastic production frontier can be written as:

$$\ln y_i = \alpha + \beta' X_i + v_i - u_i, \quad (1)$$

where α is a constant, y_i represents the output of each farm i , X_i is a vector of inputs, β' is a vector of the parameters to be estimated, v_i captures statistical noise and u_i represents the inefficiency term.

According to the original model specification, maximum likelihood estimates are obtained under these assumptions (Coelli et al., 2005):

- $v_i \sim iid N(0, \sigma_v^2)$; (2)
- $u_i \sim iid N^+(0, \sigma_u^2)$. (3)

Assumption (2) means that values of v_i are independently and identically distributed normal random variables with zero means and variances σ_v^2 . Assumption (3) expresses that values of u_i are independently and identically distributed half-normal random variables with zero means and variances σ_u^2 .

In most cases, the estimation of a stochastic production frontier aims to predict technical (in)efficiency. The most common output-oriented measure of TE is the ratio of observed output to the corresponding stochastic frontier output (Coelli et al., 2005):

$$TE_i = \frac{y_i}{\exp(\beta' X_i + v_i)} = \frac{\exp(\beta' X_i + v_i - u_i)}{\exp(\beta' X_i + v_i)} = \exp(-u_i). \quad (4)$$

One disadvantage of the above model is that it assumes that every farm uses the same technology; however, there may be latent technological differences among farms.

There are two main methods for estimating different technologies. The first and most commonly used is the ‘two-step’ method. In the first step, the sample is divided based on a

priori information, and in the second step, different frontiers are applied to the obtained groups. In the second method, which is generally considered more advanced, technological differences are taken into consideration by means of econometric measures, and the potential production level is determined and the TE estimated in one step. Some of the more important models that use this second method are the RPM and the LCM (see, among others, Greene, 2005; Orea and Kumbhakar, 2003; del Corral et al., 2009; Alvarez and del Corral, 2010). A disadvantage of the two-step method is that a single exogenous a priori piece of information cannot accurately estimate all of the technological differences between farms. In contrast, the one-step methods consider all of the information contained in the data (del Corral et al., 2009).

In this article we use an LCM. LCMs assume that there are a finite number of groups underlying the data and estimate a different frontier for each of these groups (Alvarez et al., 2012).

Following Alvarez and del Corral (2010), an LCM might be written as follows:

$$y_{it} = f(\mathbf{X}_{it})|j * \exp(v_{it}|j - u_{it}|j), \quad (5)$$

where i represents farms, t denotes time and j indicates the different classes. The vertical bar means that there is a different model for each class j .

For the half-normal model, the likelihood function (LF) for each farm i at time t for group j can be written as (Kumbhakar and Lovell, 2000):

$$LF_{ijt} = f(y_{it}|\mathbf{X}_{it}, \beta_j, \sigma_j, \lambda_j) = \frac{\Phi(-\lambda_j * \varepsilon_{it}|j / \sigma_j)}{\Phi(0)} * \frac{1}{\sigma_j} * \phi\left(\frac{\varepsilon_{it}|j}{\sigma_j}\right), \quad (6)$$

where $\varepsilon_{it}|j = y_{it} - \beta'_j x_{it}$, $\sigma_j = [\sigma_{uj}^2 + \sigma_{vj}^2]^{1/2}$, $\lambda_j = \sigma_{uj}/\sigma_{vj}$ and ϕ and Φ denote the standard normal density and cumulative distribution function, respectively.

The contribution of farm i to the conditional (on class j) likelihood is (Greene, 2005):

$$LF_{ij} = \prod_{t=1}^T LF_{ijt}. \quad (7)$$

The LF for each farm is obtained as a weighted average of its LF for each group j , using the prior probabilities of class j membership as weights (Alvarez and del Corral, 2010):

$$LF_i = \sum_{j=1}^J P_{ij} LF_{ij}. \quad (8)$$

There are many ways to parameterise P_{ij} (Greene, 2005); a convenient way is the multinomial logit (Greene, 2005, Alvarez and del Corral, 2010):

$$P_{ij} = \frac{\exp(\delta_j q_i)}{\sum_{j=1}^J \exp(\delta_j q_i)}, \quad (9)$$

where q_i is a vector of the separating variables, which are individual (farm) characteristics that sharpen the prior probabilities, and δ_j is a vector of the parameters to be estimated.

The equation of the overall log LF is the sum of the individual log LFs (Greene, 2005; Alvarez and del Corral, 2010):

$$\log LF = \sum_{i=1}^N \log LF_i. \quad (10)$$

Given that LF is maximised with respect to the parameter set using conventional methods, the estimated parameters can be used to estimate the posterior probabilities of class membership using the Bayes theorem (Alvarez and del Corral, 2010):

$$P(ji) = \frac{P_{ij} LF_{ij}}{\sum_{j=1}^J P_{ij} LF_{ij}} \quad (11)$$

In standard SFA, where the frontier function is the same for every firm, (in)efficiency is estimated relative to the frontier for all observations (see equation 4). However, in an LCM, (in)efficiency is estimated relative to as many frontiers as the number of classes. The TE can be computed using the posterior probability for each farm, and assign it a class based on the highest probability. Once the class assignment is complete, the inefficiency for that firm is computed (according to equation 4) using the frontier assigned for that class as its reference technology (Orea and Kumbhakar, 2003, p. 7).

Studies revealed in the dairy sector, that the most important factor which differentiates farms into unique technologies is intensity of production. Furthermore, Alvarez-Del Corral (2010) report that the effect of intensification may partially be confounded with size. Hence, we

assume that in the crop sector, land size might be a good proxy for both scale and intensification of production and thus we use it as a separation variable in the estimated LCM.

The translog functional form was used in the estimation and the LIMDEP software package was used for estimation purposes.

The empirical model can be written as follows:

$$\ln y_{it} = \beta_0|_j + \sum_{k=1}^K \beta_k|_j \ln X_{kit} + \frac{1}{2} \sum_{k=1}^K \sum_{l=1}^K \beta_{kl}|_j \ln X_{kit} \ln X_{lit} + \beta_t|_j * t + \frac{1}{2} \beta_{tt}|_j t^2 + \beta_{kt}|_j \ln X_{kit} * t + v_{it}|_j - u_{it}|_j \quad (12)$$

, where t is a time trend representing technical change.

5. Data

For the empirical analysis, we used data from the Hungarian Farm Accountancy Data Network (FADN). The Hungarian FADN system contains data from about 1900 annually reporting agricultural farms, and the database is representative of the farm population of over two European Size Units (ESU) (Keszthelyi, 2010). According to the EU FADN typology, ESU is used to express the economic size of a farm in the EU. It is calculated by the SGM¹ value expressed in euros divided by 1200 (EC, 1985). The agricultural farms were selected using the Hungarian Central Statistical Office's general agricultural census and Farm Structure Survey by means of a stratified sampling method (Kapronczai, 2007). The farms were categorised into different subsamples according to farm size, type of farming, legal form and other factors.

We used data on specialised COP farms over the 2001–2009 periods. Type of farming in the EU FADN typology is defined in terms of the relative importance of a specific part of a farm's total farming activity, that is, a certain crop or animal category. Relative importance is measured quantitatively as a proportion of each enterprise's standard gross margin (SGM) to the farm's total SGM (EC, 1985). According to the EU FADN typology, a farm can be defined as a specialised COP producing farm if the SGMs of cereal oilseeds and protein crops are greater than or equal to two thirds of the farm's total SGM (EC, 1985). Agricultural farms can join

¹ The standard gross margin (SGM) is the average value of output minus certain specific costs of each agricultural product (crop or livestock) in a given region. To avoid bias caused by fluctuations, such as in production (due to bad weather) or in input/output prices, basic data for a reference period of three successive years are used by Member States to calculate the SMG coefficients. The SGM of the holding is calculated as the sum of the SGM of each agricultural product present in the holding, multiplied by the relevant number of its hectares or heads of livestock (EC, 1985).

and leave the Hungarian FADN system, and to maintain representativeness, farms that leave the system are replaced by similarly characterised farms (Keszthelyi and Pesti, 2009). Our primary goal in this study was to explore the latent technological differences between farms and investigate the effects of these differences in TE. This question can be better examined if farms with the same, and not just similar, characteristics are observed, which is why we used balanced panel data (i.e., we used only those farms that were observable every year). Our sample contained 1656 observations, 184 for each year. The data were provided by the Research Institute for Agricultural Economics.

For the purpose of estimation, one output (Y – total agricultural production in value) and four inputs (labour in Annual Work Units² [X_1], utilised agricultural area (UAA) in hectares [X_2], total fixed assets in value [X_3] and total specific costs in value [X_4]) were used. Total specific costs were equal to crop-specific (seeds and seedlings, fertilisers, crop protection products, other specific crop costs) and livestock-specific (feed for grazing stock and granivores, other specific livestock costs) inputs (EC, 2007).

Additionally, a time variable (t) and time-squared variable (tt) were added to the production frontier to allow for non-monotonic technical change (i.e., allowing the TCH effect to increase ($\beta_{tt}>0$) or decrease ($\beta_{tt}<0$)), and the time trend was interacted with the input variables to allow for non-neutral technical change³ (i.e., allowing the marginal rates of technical substitution to change). All of the variables expressed in nominal prices were deflated to 2005 prices with the use of the appropriate deflators; precisely, the output (Y) was deflated by the agricultural output price index, the total specific costs (X_4) by the price index of purchased goods and services and the corresponding values of total fixed assets (X_3) by the price index of agricultural investments.

Some of the descriptive statistics for the sample are included in Table 1. The high variance of the individual variables was quite apparent. The labour input had a minimum value of 0.01 AWU and a maximum of 87 AWU, and the values for (UAA) ranged from 8.5 to 3837 hectares.

Table 1: Variable definitions and descriptive statistics

² One Annual Work Unit (AWU) corresponds to the work performed by one person on an agricultural farm performing agricultural activities on a full-time basis (1 AWU = 1800 working hours = 225 working days. (Keszthelyi and Pesti, 2009).

³ If the marginal rate of technical substitution is independent of time, then technical change is said to be Hicks neutral (Coelli et al., 2005, p. 213).

Variables	Symbol	Mean	Standard deviation	Minimum	Maximum
Total output (1000 HUF)	Y	34879.1	76877.7	321.009	916979
Labour (AWU)	X ₁	3.69657	8.79723	0.01	86.79
UAA (ha)	X ₂	245.133	476.009	8.5	3836.9
Total fixed assets (1000 HUF)	X ₃	46037.3	62969.3	240.001	555270
Total specific costs (1000 HUF)	X ₄	12454.6	28089.2	56.1101	273696

Source: Authors' calculations based on Hungarian FADN data.

These remarkable variances in input values indicate inequalities in the farm structure, and these inequalities are well illustrated in Figure 2, which shows the Lorenz curves of the individual inputs. The Lorenz curves demonstrate significant inequalities for all of the inputs. Both the descriptive statistics and the Lorenz curves suggest that farm size probably plays a role in the technological differences between farms. Therefore, in the course of the empirical estimation of the model, we used the size of the agricultural area as the separating variable to sharpen the probability of belonging to a given group. Although the difference in the case of the other inputs was also remarkable, to avoid correlation among the separation variables we used only the agricultural area.

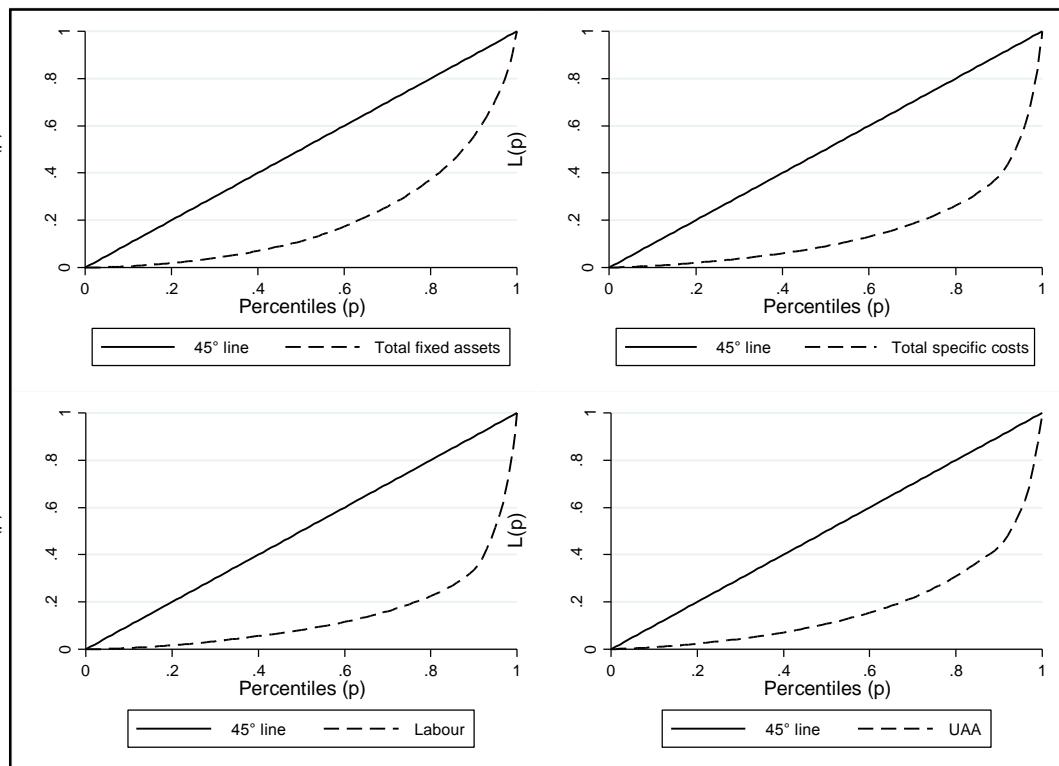


Figure 2: Lorenz curves of inputs
Source: Authors' illustration based on Hungarian FADN data.

6. Econometric results

6.1. Parameter estimates

In LCMs, the number of groups must be specified, as it is not a parameter to be estimated (Alvarez and del Corral, 2010). We tried to estimate the model with different numbers of classes. The model with three classes failed to achieve convergence and, following Orea and Kumbhakar (2003), we take this as evidence that a model with three classes is overspecified. The model with two classes was estimable in Table 2, the parameter estimates of the LCM with the obtained two classes (C1 and C2) are compared with the results of a model that assumes a common frontier to all farms⁴. The number of farms in class 1 and class 2 is 855 and 805, respectively. The input variables were normalised with their geometric means; consequently, the obtained parameters can be considered as output elasticities evaluated at the geometric mean of the sample.

With regard to both models and both groups, the signs of the elasticity of labour, agricultural area, fixed assets and specific costs met our expectations; that is, all of them were positive.

The sum of the output elasticities is 1,03 one in the case of the common frontier model and near one in the case of the LCM indicating slightly decreasing return to scale and constant return to scale, respectively, implying that farms have exploited the full potential of the technology they are currently working.

The value of the parameter λ was significant in both models. This parameter provides information about the importance of an inefficiency effect compared to statistical noise. If λ is greater than 1, the deviation from the theoretic production frontier is mostly due to the inefficiency effect. In the present case λ is greater than one in both models, meaning that technical inefficiency is an important phenomenon in Hungarian agriculture and must be included in the production model. Moreover, this suggests that there is some room to increase productivity through TE in the Hungarian agriculture. Thus, policies focusing on agricultural education and extension programs might be effective measures.

⁴ We note that in the text and in the tables the terms "group" and "class" denote the same.

The models showed positive TCH for the period analysed ($\beta_t > 0$), but the rate of development was decreasing ($\beta_{tt} < 0$). Technological development was labour saving ($\beta_{tx1} < 0$) and specific cost using ($\beta_{tx4} > 0$) based on the results of the model.

Table 2: Parameter estimates^a

	Common frontier		LCM-Group 1 (C1) ^b		LCM-Group 2 (C2) ^b	
	Coefficients	t- ratio	Coefficients	t- ratio	Coefficients	t- ratio
Constant	0.33036***	13.2	0.25257***	8.5	0.48207***	17.7
t	0.01786***	4.5	0.02449***	5.0	0.01642***	3.8
tt	-0.00957***	-2.9	-0.01573***	-3.8	-0.00750**	-2.0
X₁	0.12997***	6.5	0.06414***	3.3	0.14136***	5.5
X₂	0.50988***	14.6	0.39412***	8.8	0.47801***	12.4
X₃	0.07639***	4.7	0.06123***	2.9	0.10907***	7.2
X₄	0.32139***	10.6	0.48804***	13.2	0.27105***	9.0
X₁*X₂	0.05449	1.1	0.02436	0.5	-0.02318	-0.3
X₁*X₃	0.01453	0.5	0.00571	0.3	0.06742**	2.2
X₁*X₄	-0.06309	-1.4	-0.0162	-0.5	-0.18168**	-2.6
X₂*X₃	.07417*	1.8	0.12288**	2.2	0.0359	0.8
X₂*X₄	0.11118	1.6	0.12801	1.2	0.24532**	2.6
X₃*X₄	-0.07254**	-2.1	-0.07564*	-1.7	-0.10647***	-3.4
X₁*X₁	0.08643**	2.3	0.06473**	2.1	0.25410***	5.0
X₂*X₂	-0.37953***	-3.7	-0.48070***	-2.9	-0.39787***	-2.8
X₃*X₃	0.00249	0.1	-0.04576**	-2.1	0.01957	0.8
X₄*X₄	0.06284	0.8	0.08465	1.0	0.04418	0.4
t*X₁	-0.02011***	-2.8	-0.02482***	-3.8	-0.02351***	-3.0
t*X₂	0.0067	0.6	-0.00796	-0.6	0.00888	0.8
t*X₃	0.00209	0.4	0.01267*	1.9	-0.00303	-0.6
t*X₄	0.01173	1.0	0.02083*	1.8	0.01864*	1.7
λ	1.24390***	8.4	2.49515***	28.1	2.04168***	21.5
σ_u	0.38386***	9.8	0.470222***	9.4	0.359649***	7.6
Log LF	-555.88		-465.67			
Akaike	0.699		0.622			
Information Criterion (AIC)						
Estimated probability for class membership						
Constant	-	-	0.53491**	2.2	-	-
UAA	-	-	-0.00202**	-2.3	-	-

^aNote: ***, ** and * denote significance at the 1%, 5% and 10% levels, respectively.

^b Group 1 (C1) and Group 2 (C2) denote the two groups obtained by the LCM.

The obtained results provide comparisons from two perspectives: the results of the common frontier with the groups obtained by the LCM and the differences between the two groups as defined by the LCM. The technological parameters were different in both comparisons. To determine which model to choose, Akaike Information Criterion (AIC) values were used. According to the literature (e.g., Alvarez and del Corral, 2010), a model with a lower AIC value fits the sample better. The corresponding values for the LCM were 0.7 (AIC), and in the case of the model with the assumption of equal technology, 0.62. This indicates that the LCM fits the sample data better.

The coefficient of the UAA used as a separation variable was significant, which proves our hypothesis that farm size plays an important role in the establishment of the two groups; that is, the increase in agricultural area reduces the probability of a farm ending up in Group 1.

6.2. Differences between the groups obtained by LCM

Our second goal was to determine whether there were indeed technological differences between the two groups obtained by the model. We compared the differences between the groups in terms of marginal products and elasticity of scale following Alvarez and del Corral (2010). Among the inputs, the marginal products of both the agricultural area and the specific costs were analysed, because they showed the most significant difference between the two groups. For the marginal product calculation, the elasticity values were multiplied by the average products of the individual farms, and then the arithmetic mean of the values obtained for the investigated period (Table 3) was taken.

	LCM		Common frontier	
	C1	C2	C1	C2
UAA	35.4	70.4	53.5	40.6
Total specific cost	3139.9	5805.1	1959.4	8656.7

Source: Authors' calculation.

To better demonstrate the differences, the marginal products in the case of the common frontier were also calculated using the LCM groups, as in Alvarez and del Corral (2010), allowing the comparison to show how the technological characteristics differed between models due only to differences in the estimated parameters (as the farms being compared

were the same). Table 3 shows that the marginal products were different between the LCM groups. In the case of Group 2 (C2) the marginal products of land and total specific cost were much higher in both models.

In terms of elasticity of scale, although less conspicuous, differences were still observable between the two groups (Table 4, Figure 3). The scale elasticity (the sum of the individual input elasticities) was calculated for each observation and the arithmetic mean of this calculation is shown in Table 4.

Table 4: Scale elasticities

LCM		Common frontier	
C1	C2	C1	C2
0.994	1.007	1.037	1.038

Source: Authors' calculations.

We used statistical tests to examine the differences between the scale elasticities of the LCM groups (see Appendix 1). Applying the Mann–Whitney test revealed a significant difference between the two groups. We used Kolgomorow–Smirnov and Kruskall–Wallis tests to determine whether the distribution of the two groups was the same and whether they could be considered homogeneous. The null hypothesis was rejected in both of the tests, thus the two groups were not homogeneous, and their distribution was not the same.

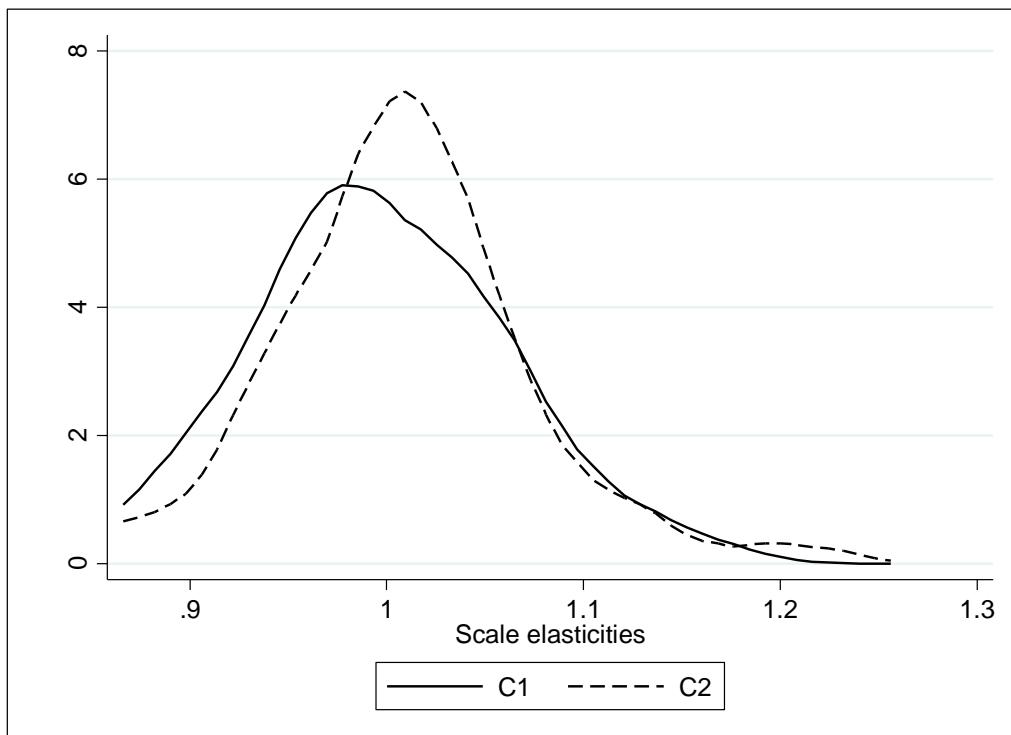


Figure 3: Kernel distributions of scale elasticities

Source: Authors' illustration.

6.3. Characteristics of the LCM groups

For a more detailed examination of the differences, certain characteristics of the groups were compared, as shown in Table 5. To examine the differences between the groups, the following characteristics were examined: size, organisational form, cost of external factors, share of irrigated land and intensity of input usage. The differences in size were compared using the output of farms, the UAA and the ESU. The differences in organisational form were analysed using the definition from Hill (1993) in which a farm is labelled as a family farm if the ratio of family work unit per annual work unit (FWU/AWU) is greater than 0.95. In the FADN database, the definition of external factors is remuneration of inputs (work, land and capital) that are not the property of the holder (wages, rent and interest paid). Finally, the differences in intensity were investigated using the ratio of the costs of fertiliser, crop protection and seed per hectare. We concluded from the data in Table 5 that the differences between the groups (except for share of irrigated land) were statistically significant, according to the Mann–Whitney test.

Table 5: Characteristics of the LCM groups

	C1	C2	Mann–Whitney test		
				z-stat.	p-value
Output (1000 HUF)	13686.4	57500.4	-14.0	0.000	
UAA (ha)	130.2	367.8	-6.5	0.000	
Economic size (ESU)	28.1	91.4	-9.3	0.000	
Share of family farms (%)	83	66	-8.2	0.000	
Share of irrigated land (%)	0.10	0.61	-1.30	0.1932	
Cost of external factors (1000 HUF/ha)	11.31	24.025	-8.34	0.000	
Fertiliser cost/ha (1000 HUF/ha)	14.9	17.3	-6.5	0,000	
Seed cost/ha (1000 HUF/ha)	12.6	15.6	-9.5	0,000	

Crop protection cost/ha (1000	10.6	13.9	-10.6	0,000
HUF/ha)				

Source: Authors' calculation.

Note: The variables in the table are averages, except for share of family farms and share of irrigated land.

Hungarian agriculture, as seen previously, is exceedingly heterogeneous regarding the economic size of farms. Our results, in line with the parameter estimates on agricultural area used as a separating variable, verified that the presumption derived from the descriptive statistics was true; that is, in the case of such considerable heterogeneity of farm size, differences in applied technologies exist and farm size is one of the major factors determining the choice of technology.

Although size was an important differentiating factor, the differences between the examined groups did not arise solely from the inequalities in farm size; that is, the two groups could not be classified along the lines of the traditional categorisation of small- and large-scale farms.

The farms also differed in their use of external factors and in production intensity. Group 2 farms used more external factors in their production, had more hired land and used more hired labour and more credit. Fertiliser, crop protection and seed costs per area unit were also higher in Group 2, which indicated that the farms in this category undertook more intensive farming activities.

Consequently, the estimation of one common frontier is not likely to represent the 'real' technological level and may result in biased TE estimates. In addition, the estimation may overestimate the level of agricultural production potential, which could lead to inadequate agricultural policy implications.

6.4 Comparison of the TE scores with a standard SFA model

The differences in the extent of TE for the models are summarised in Table 6. As expected, according to the literature (see, e.g., Alvarez et al., 2012, Alvarez and del Corral, 2010; Sauer and Morrison, 2013), the TE of both groups was higher with the LCM than with the model assuming the same technology. The reason for this is that the LCM allows each farm's TE to

be measured with respect to its own frontier (Alvarez and del Corral, 2010). Group 2 showed a higher degree of TE in both models, but with the LCM the difference between the two groups was less significant. The usefulness of the LCM can be well seen here. The result of the estimated common frontier suggested that farms belonging to Group 1 could increase their output by 40% while leaving input levels unchanged, whereas the LCM showed a 15% increase. We also observed significant differences in the average values of TE, with the value from the LC model being almost 10% higher.

Table 6: TE averages for C1, C2 and for the whole sample

LCM			Common frontier		
C1	C2	whole sample	C1	C2	whole sample
0.810	0.857	0.83	0.637	0.847	0.74

Source: Authors' calculations.

6.5 Productivity differences between the identified technologies

As the final step of our analysis, we examined the productivity differences between groups. First, we compare the predicted outputs of the estimated frontiers (Y_1 and Y_2) and then the average yields of the main products of the farms. Y_1 represents the predicted output of the C1 frontier, while Y_2 denotes the predicted output of the C2 frontier. The estimated outputs are calculated based on equation 12. More precisely, we calculate both Y_1 and Y_2 for every farm using the input quantity of a given farm and first, assuming (a), that the farm uses C1 technology (i.e. the calculation is made using the parameter estimates of the C1 frontier) and second (b), that the farm uses C2 technology (i.e. the calculation is made using the parameter estimates of the C2 frontier). Y_1 and Y_2 are always calculated for every farm, irrespective of the class the farm is assigned by the LCM.

According to Alvarez and del Corral (2010) and Kumbhakar et al., (2009), if the predicted output is higher for technology 'a' than for technology 'b', then, assuming a given input level, technology 'a' is locally above technology 'b'. We can conclude that one technology is more productive than the other only if all data points for this technology are above those of the other. Table 7 contains the averages, variances and minimum and maximum values of the predicted outputs in the years under study. On the basis of the data in the table, we concluded that the technology of Group 2 to be the more productive. The predicted average output of this

group is higher than that of the other group, and the average difference was more than 30% between groups. We find that Group 2 frontier is located above the Group 1 frontier for almost 98% of the cases (see Appendix 2). Although the abovementioned criterion was not satisfied 100%, it seems to be that the productivity of Group 2 is higher, particularly when bearing in mind that the TE of Group 2 is also higher and the TE difference between the groups is statistically significant (see Appendix 3)

Table 7: Comparison of predicted outputs

		Means	Std. dev.	Minimum	Maximum
2001	Y2	40730.7	85700.5	2148.62	667335
	Y1	32448.1	87148.2	1428.31	956473
	Y2/Y1	1.42	0.29	0.70	3.11
2002	Y2	42641.6	88934	2389.78	676116
	Y1	36157.7	93983.3	1846.76	931922
	Y2/Y1	1.37	0.23	0.73	2.82
2003	Y2	46753.5	95399.5	2468.54	731301
	Y1	39332.9	97817.1	1904.27	959923
	Y2/Y1	1.37	0.18	0.76	2.15
2004	Y2	49965.7	102915	2469.86	902294
	Y1	42619.6	105684	1675.22	1.06E+06
	Y2/Y1	1.34	0.19	0.85	2.25
2005	Y2	50550.9	99217.5	2638.01	877362
	Y1	42077.5	102784	2066.59	1.12E+06
	Y2/Y1	1.33	0.20	0.79	2.57
2006	Y2	51805.4	101623	2065.03	882330
	Y1	43040.3	98028.3	1570.82	975014
	Y2/Y1	1.31	0.17	0.90	2.17
2007	Y2	52831.8	102970	1879.92	851156
	Y1	44874.4	102177	1170.69	950415
	Y2/Y1	1.32	0.16	0.90	2.06
2008	Y2	53080.9	101768	1045.59	774926
	Y1	45380.6	103720	1121.42	888778
	Y2/Y1	1.33	0.20	0.85	2.60
2009	Y2	53047.3	99786	1718.43	747711
	Y1	44431	95102.4	1098.59	751232
	Y2/Y1	1.34	0.18	0.89	2.07

Source: Authors' calculations.

A comparison of average yields reinforced the conclusion drawn from the predicted Y values that the productivity of the Group 2 farms was higher than that of the Group 1 farms (Table 8). The average yields of the Group 2 farms showed similarly higher values for all products, the difference being smallest for winter barley and largest for maize.

Table 8: Average crop yields

	C1	C2	C2/C1
Winter wheat	3.24	3.97	1.23
Winter barley	0.91	1.07	1.18
Maize	3.98	6.29	1.58
Rapeseed	0.49	0.59	1.20
Sunflowers	1.24	1.66	1.34

Source: Authors' calculations.

7. Conclusions and policy implications

In this paper, we analyse the TE gap in the context of Hungarian COP farms and use an LCM to examine the effect of technological heterogeneity on the estimated TE scores. Our analysis has a number of interesting methodological and agricultural policy implications.

From a methodological point of view, first we provide evidence in favour of our hypotheses that it is especially important to account for technological differences when examining transition countries and COP farms, which are generally considered more homogeneous. Second, the results of our analyses reveal that the TE in the COP farming sector is higher than would be expected from the results of earlier studies using traditional models without controlling for technological heterogeneity. Third, using an LCM, we identify two classes of Hungarian COP producing farms that use different technologies and find that the scale of land use is one of the most important factors that differentiates our sample into unique technologies.

From an agricultural policy point of view, our finding that TE appears to be higher in the COP sector indicates that there is less opportunity to increase performance through TE

improvement than was earlier expected. Moreover, our results highlight that both of the identified groups of farms appear to be working near to the constant return to scale part of the production frontier, implying that the farms have exploited the full potential of the technology they are currently using. This also suggests that there is little chance to improve farm performance by increasing in size, unless the farms switch technology. Consequently, agricultural policies designed to increase productivity in the Hungarian COP sector should concentrate on technological progress.

As one of the central aims of the CAP is to increase productivity, our results are also likely to be useful in the context of the CAP. The new agreement on CAP reform aims to enhance the effectiveness of the CAP through the better targeting of agricultural support by offering member states the opportunity to target direct payments through optional support schemes. Therefore, our results may help decision makers who want to use these optional support schemes to increase productivity in the following ways. The finding that TE is higher if we take into account the effect of different technologies is in line with papers analysed the dairy sector in different countries (Alvarez-Del Corral, 2010; Alvarez et al., 2012; Emvalomatis, 2007; Sauer-Morrison, 2013). Thus, this finding appears to be general, irrespective of agricultural sectors or countries. Consequently, our results suggest that policy makers should select R&D as a policy option, as this would have a greater influence on productivity improvement than directing support schemes towards extension.

Acknowledgements

Lajos Baráth would like to acknowledge the support of the OTKA 0038 programme and the help of the Swiss programme: “The impact of insurance on the economic performance of Hungarian cropping farms”.

The authors would like to acknowledge the Research Institute for Agricultural Economics for providing the FADN data.

Appendix 1: Examination of Scale Elasticity (SE) differences using statistical tests between the obtained LCM classes (C1 and C2)

1. Two-sample Wilcoxon rank-sum (Mann–Whitney) test

Classes	obs	rank sum	expected
C1	963	751920.5	797845.5
C2	693	620075.5	574150.5

H_0 : The null hypothesis of this test is that the SE values in the two groups (C1 and C2) are from populations with the same distribution.

$z = -4.784$

$\text{Prob} > |z| = 0.0000$

2. Two-sample Kolmogorov–Smirnov test for equality of distribution functions

Classes	D	P-value	Corrected
C1	0.2871	0.000	-
C2	-0.2043	0.000	-
Combined K-S:	0.2871	0.000	0.000

H_0 : The null hypothesis of this test is that the distribution of SE in Group 1 and Group2 are the same.

3. Kruskal–Wallis equality-of-populations rank test

Classes	Obs	Rank Sum
C1	963	751920.5
C2	693	620075.5

H_0 : The null hypothesis of this test is that the SE values are from the same population.

Chi-square = 22.888 with 1 d.f.

Probability = 0.0001

Appendix 2: Distribution of the predicted output ratios (Y2 versus Y1)

Y2/Y1 ratio	Number of observations	Percentage of observations
Y2/Y1 ratio<=1	36	2.17
1<Y2/Y1 ratio<=1.2	214	12.92
1.2<Y2/Y1 ratio<=1.3	449	27.11
1.3<Y2/Y1 ratio<=1.4	508	30.68

1.4<Y2/Y1 ratio<=1.5	232	14.01
Y2/Y1 ratio>1.5	217	13.10

Appendix 3: Examination of TE differences using statistical tests between the obtained LCM classes (C1 and C2)

Classes	obs	rank sum	expected
C1	855	612808	708367.5
C2	801	759188	663628.5
combined	1656	1371996	1371996

H_0 : The null hypothesis of this test is that the TE values in the two groups (C1 and C2) are from populations with the same distribution.

$z = -9.827$

Prob > |z| = 0.0000

References

- Aigner, D., Lovell, C., Schimdt, P. 1977. Formulation and estimation of stochastic production function models. *Journal of Econometrics*. 6, 21-37.
- Alvarez, A., Arias, A., Greene, W. 2003. Fixed management and time invariant technical efficiency in a random coefficient model. Working Paper. Department of Economics, Stern School of Business, New York University. 22 pp.
- Alvarez, A., del Corral, J. 2010. Identifying different technologies using a latent class model: extensive versus intensive dairy farms. *European Review of Agricultural Economics*. 37 (2), 231-250.
- Alvarez, A., del Corral, J., Tauer, L.W. 2012. Modeling unobserved heterogeneity in New York dairy farms: One-stage versus two-stage models. *Agricultural and Resource Economics Review*. 41 (3), 275-285.
- Bakucs L.Z., Latruffe, L., Fertő, I., Fogarasi, J. 2010. The impact of EU accession on farms' technical efficiency in Hungary. *Post-Communist Economies*. 22 (2), 165-175.
- Bakucs, L.Z., Fertő I., Fogarasi J., Tóth J. 2012. Farm organisation and efficiency in Hungarian dairy farms. *Milk Science International*. 67 (2), 147-150.
- Banse, M., Gorton, M., Hartel, J., Hughes, G., Kockler, J., Mollman, T., Munch, W. 1999. The evolution of competitiveness in Hungarian agriculture: from transition to accession. *MOCT-MOST*. 9, 307-318.
- Bignebat, c., Latruffe, L. 2009. Twenty years of land reforms in central and eastern Europe: state of play and outlook. Working Paper. SMART-LEREKO Nr. 09-19. 20.
- Bojnec, Š., Latruffe, L. 2013. Farm size, agricultural subsidies and farm performance in Slovenia. *Land Use Policy*. 32, 207-217.

- Cechura, L. 2010. Estimation of technical efficiency in Czech agriculture with respect to firm heterogeneity. *Agricultural Economics*. 56 (4), 183-191.
- Ciaian, P., Pokrivicak, J., Drabik, D. 2009. Transaction costs, product specialisation and farm structure in Central and Eastern Europe. *Post Economist Economies*. 21(2), 191-201.
- Coelli, T.J., Rao, D.S.P., O'Donnell, C.J., Battese, G.E. 2005. *An Introduction to Efficiency and Productivity Analysis*. Springer, USA, 347 pp.
- Csáki, Cs., Lerman, Z. 1999. *Structural Change in the Farming Sectors in Central and Eastern Europe*. World Bank, Washington D.C., 243 pp.
- CSO. 2012. Agricultural Census 2010. KSH, Budapest. 34 pp.
- del Corral, J., Alvarez, A., Tauer, L. 2009. Detecting technological heterogeneity in New York farms. Working Paper. Department of Applied Economics and Management. Cornell University, Ithaca, New York 14853-7801 USA. WP 2009-16, April 2009. 11 pp.
- EC (1985). Commission Decision of 7 June 1985 establishing a Community typology for agricultural holdings (85/377/EEC). *Official Journal of the European Communities*. No L 220/1- 32.
- EC (2007). Definitions of variables used in FADN standard results. RI/CC 882 Rev.8.1. European Commission, Brussels. 12 April 2007. 32 pp.
- EC (2013). Overview of CAP reform 2014-2020. *Agricultural Policy Perspectives Brief*. Nr. 5, December 2013. European Commission, Brussels, 10 pp.
- Emvalomatis, G. 2007. Distinguishing technologies and measuring performance in farm production: The case of Pennsylvania dairy. A thesis in Agricultural, Environmental and regional Economics. The Pennsylvania State University, August 2007. 56 pp.
- Fogarasi, J., Latrufe, L. 2009. Technical efficiency in dairy farming: a comparison of France and Hungary in 2001-2006. *Studies in Agriculture Economics*. 110, 75-84.
- Gorton, M., Davidova, S. 2004. Farm productivity and efficiency in the CEE applicant countries: a synthesis of results. *Agricultural Economics*. 30, 1-16.
- Greene, W. 2005. Reconsidering heterogeneity in panel data estimators of the stochastic frontier model. *Journal of Econometrics*. 126, 269-303.
- Harangi-Rákosi, M., Szabó, G. 2011. A mezőgazdasági szervezetek gazdálkodásának vizsgálata a 2002–2009 közötti időszakban (in Hungarian with English abstract). *Gazdálkodás*. 4, 358-366.
- Hill, B. 1993. The 'Myth' of the Family Farm: Defining the Family Farm and Assessing its Importance in the European Community. *Journal of Rural Studies*. 9 (4). 359–370.
- Hockmann, H. , Pieniadz, A. 2007. Farm heterogeneity and efficiency in Polish agriculture: a stochastic frontier analysis. 104th (joint) EAAE-IAAE Seminar Agricultural Economics and Transition: "What was expected, what was observed, the lessons learned" Corvinus University of Budapest (CUB) Budapest, Hungary, September 6-8, 2007.

Jámbor, A. 2009. A magyar gabonafélék és feldolgozott termékeinek komparatív előnyei és versenyképessége az EU-15 országok piacain (in Hungarian with English abstract). Közgazdasági Szemle. 5. 443-463.

Kapronczai, I. 2007. A mezőgazdasági jövedelem információs rendszerek összefüggései (in Hungarian with English abstract). Agrárgazdasági Információk. 1. AKI, Budapest, 120 pp.

Kapronczai, I. 2010. A magyar agrárgazdaság az adatok tükrében ez EU csatlakozás tükrében (in Hungarian with English abstract). Agrárgazdasági Információk. 12. AKI, Budapest. 152 pp.

Keszthelyi, Sz. 2010. A Tesztüzemi Információs Rendszer 2006. évi eredményei (in Hungarian with English abstract). Agrárgazdasági Információk. 5. AKI, Budapest. 39 pp.

Keszthelyi Sz., Pesti Cs. 2009. A Tesztüzemi Információs rendszer 2008. évi eredményei (in Hungarian with English abstract). Agrárgazdasági Információk. 3. AKI, Budapest. 45 pp.

Kumbhakar, S., Tsionas, E., Sipiläinen, T. 2009. Joint estimation of technology choice and technical efficiency: an application to organic and conventional dairy farming. Journal of Productivity Analysis. 31, 151-161.

Kumbhakar. S.C., Knox Lovell, C.A. 2000. Stochastic Frontier Analysis. Cambridge University Press, Cambridge, 339 pp.

Latruffe, L., Balcombe, K.G., Davidova, S., Zawalinska, K. 2004. Determinants of technical efficiency of crop and livestock farms in Poland. Applied Economics. 36 (12),1255-1263.

Meeusen, W., van den Broeck, J. 1977. Efficiency estimation from Cobb–Douglas production functions with composed error. International Economic Review. 18, 435-444.

Orea, L., Kumbhakar, S.C. 2003. Efficiency measurement using a latent class stochastic frontier model. Working Paper. http://bingweb.binghamton.edu/~kkar/ee_0251.pdf.

Pesti Cs., Keszthelyi Sz. 2010. A különböző típusú üzemek jövedelmét befolyásoló tényezők vizsgálata és nemzetközi összehasonlítása (in Hungarian with English abstract). Agrárgazdasági Információk. 2. AKI, Budapest, 63 pp.

Sauer, J., Morrison Paul, C.J. 2013. The empirical identification of heterogeneous technologies and technical change. Applied Economics. 45 (11), 1461-1479.