Farm Heterogeneity and Efficiency in Polish Agriculture: A Stochastic Frontier Analysis

HEINRICH HOCKMANN, AGATA PIENIADZ

Leibniz Institute of Agricultural Development in Central and Eastern Europe (IAMO), Theodor-Lieser-Str. 2, 06120 Halle (Saale), Germany

Email: hockmann@iamo.de.

Paper prepared for presentation at the 104th (joint) EAAE-IAAE Seminar Agricultural Economics and Transition:

„What was expected, what we observed,

the lessons learned."

Corvinus University of Budapest (CUB)

Budapest, Hungary. September 6-8, 2007

FARM HETEROGENEITY AND EFFICIENCY IN POLISH AGRICULTURE: A STOCHASTIC FRONTIER ANALYSIS

HEINRICH HOCKMANN, AGATA PIENIADZ *

ABSTRACT

This paper deals with the estimation of a random coefficient model. The virtue of this approach is that it considers firm heterogeneity, which conventional SFA models do not. Applying the model to Polish farms, the results indicate that the conventional random and fixed effect models overestimate the inefficiency score. In addition, the reasons for inefficiency are analysed. It is shown that despite the fragmentation of Polish agriculture, there is no evidence for scale inefficiency. Moreover, inefficiency could partly be attributed to factors, which affect the management input and requirements on farms.

Keywords: SFA, random component model, Poland, agriculture, management

1 INTRODUCTION

There are numerous technical and economic efficiency analyses of agriculture in central and eastern European countries (CEECs). Further, nonparametric but deterministic approaches (DEA), as well as stochastic but parametric approaches (SFA) have been widely applied (see for instance BACKUS et al. 2006; BRÜMMER et al. 2002; MUNROE, et al. 2001; LATRUFFE et al. 2004). SFA and DEA assume that farms are not heterogeneous but inefficient, since all inefficiency scores are estimated by assuming a homogeneous technology available to all producers. This again suggests that the impact of inefficiency in the agriculture of CEECs is overestimated, and, in addition, that the reasons for inefficiency might not be well identified.

We use a random coefficient specification of production technology that avoids the heterogeneity bias. Further, we follow an approach developed by Alvarez et al., (2003, 2004). Our empirical application deals with Polish agriculture, which is often labelled as 'backward' or 'inefficient'. Indeed, its weak economic performance is explained by high fragmentation, over-employment and the utilization of outdated technologies. These characteristics suggest the existence of multiple market failures, especially on the labour and capital market, but also on the product market. However, small-scale farming did not disappear during transition. This implies that such farms react flexibly to severe conditions on the factor and product markets. Following these developments, two basic questions arise, both of which will be addressed in our study:

(1) Are small farms less efficient than larger farms, i.e., is scale efficiency a significant issue in Polish agriculture?

(2) Which factors hamper efficient production?

* Leibniz Institute of Agricultural Development in Central and Eastern Europe (IAMO), Halle (Saale), Germany. Email: hockmann@iamo.de.

2 THEORETICAL BACKGROUND

The theoretical framework is developed within a panel data framework, with i = 1,…,N firms and t = 1,..,T observation per firm. We follow the factor multiplication approach and assume a production technology in which effective outputs ($\mathbf{y}^e_{it}$) are produced with observable input ($\mathbf{x}^e_{it}$). The augmented inputs and outputs are given by:

(1) $\mathbf{y}^e_{it} = \mathbf{y}_{it} e^{\boldsymbol{\tau}_{yt} t} e^{\boldsymbol{\mu}_{yi} m_i}$ and $\mathbf{x}^e_{it} = \mathbf{x}_{it} e^{\boldsymbol{\tau}_{xt} t} e^{\boldsymbol{\mu}_{xi} m_i}$.

Here, $\mathbf{y}_{it}$ and $\mathbf{x}_{it}$ represent observable inputs and outputs, t accounts for productivity change over time and m_i represents a non-observable firm specific factor In principle, m captures the *environment* of producing, and covers differences in factor qualities such us climate condition, soil fertility and human capital, including management skills, etc. We specify technology as a translog output distance function ($D_o(\mathbf{y}^e_{it}, \mathbf{x}^e_{it})$). Rearranging terms provides:

$$
\begin{aligned}
0 \geq \ln D_o\left(\mathbf{x}^e{}_{it}, \mathbf{y}^e{}_{it}\right) = {} & \alpha_0 + \alpha_m m_i + \frac{1}{2}\alpha_{mm} m_i^2 + (\alpha_t + \alpha_{tm} m_i)t + \frac{1}{2}\alpha_{tt} t^2 \\
& + (\boldsymbol{\alpha}_{\mathbf{x}} + \boldsymbol{\alpha}_{\mathbf{x}t} t + \boldsymbol{\alpha}_{\mathbf{x}m} m_i)' \ln \mathbf{x}_{it} + \frac{1}{2} \ln \mathbf{x}_{it}' \mathbf{A}_{\mathbf{xx}} \ln \mathbf{x}_{it} \\
& + \left(\boldsymbol{\alpha}_{\mathbf{y}} + \boldsymbol{\alpha}_{\mathbf{y}t} t + \boldsymbol{\alpha}_{\mathbf{y}m} m_i\right)' \ln \mathbf{y}_{it} + \frac{1}{2} \ln \mathbf{y}_{it}' \mathbf{A}_{\mathbf{yy}} \ln \mathbf{y}_{it} \\
& + \ln \mathbf{x}_{it}' \mathbf{A}_{\mathbf{xy}} \ln \mathbf{y}_{it}
\end{aligned} \qquad (2)
$$

The various parameters associated with t and m_i are functions of the original parameters $\boldsymbol{\alpha}_{\mathbf{x}}, \boldsymbol{\alpha}_{\mathbf{y}}, \mathbf{A}_{\mathbf{xx}}, \mathbf{A}_{\mathbf{yy}}, \mathbf{A}_{\mathbf{xy}}$ as well as the productivity terms $\boldsymbol{\tau}_{\mathbf{x}t}, \boldsymbol{\tau}_{\mathbf{y}t}, \boldsymbol{\mu}_{\mathbf{x}i}, \boldsymbol{\mu}_{\mathbf{y}i}$. Technical efficiency can be introduced by assuming, that actual m_i is not necessarily at its optimal level (m_i^*). Accordingly, we define the technical efficiency as:

(3) $\ln TE_{it} = \ln D_o\left(\mathbf{x}^e{}_{it}, \mathbf{y}^e{}_{it}\right) - \ln D_o\left(\mathbf{x}^e{}_{it}, \mathbf{y}^e{}_{it}\right)_{|m_i = m_i^*} \leq 0$.

Thus, the last inequality results from the fact that the output distance function with optimal firm-specific effects is efficient. Since neither m_i nor m_i* are observable, (3) cannot be estimated directly. ALVAREZ et al. (2003, 2004) develop an estimable model. From (2) and (3) it follows:

(4') $0 \geq \ln D_o\left(\mathbf{x}^e{}_{it}, \mathbf{y}^e{}_{it}\right)_{|m_i = m_i^*} + \ln TE_{it}$.

Considering that the output distance function is linearly homogenous in output provides:

(4) $\ln y_{it} \geq \ln D_o\left(\mathbf{x}^e{}_{it}, \tilde{\mathbf{y}}^e{}_{it}\right)_{|m_i = m_i^*} + \ln TE_{it}$,

where $\tilde{\mathbf{y}}^e_{it}$ represent normalized outputs. Equation (4) can be estimated by maximum simulated likelihood with the following distributional assumptions $\ln TE_{it} \sim N^+(0, \sigma_u)$, $m_i^* \sim \bullet(0,1)$. The symbol • indicates that m_i* might possess any distribution with zero mean and unit variance. In addition, random effects are considered in a variable $v_{it} \sim N(0, \sigma_v)$. Moreover, TE_{it} is defined by:

$$\ln TE_{it} = \gamma_0 + \gamma_t t + \boldsymbol{\gamma}_{\mathbf{x}}'\ln \mathbf{x}_{it} + \boldsymbol{\gamma}_{\mathbf{y}}'\ln \mathbf{y}_{it}, \text{ with}$$

$$(5)\qquad \begin{aligned} \gamma_0 &= \alpha_m (m_i - m_i^*) + \frac{1}{2}\alpha_{mm}\left(m_i^2 - m_i^{*2}\right) \\ \gamma_t &= \alpha_{tm}(m_i - m_i^*) \\ \boldsymbol{\gamma}_{\mathbf{x}} &= \boldsymbol{\alpha}_{\mathbf{x}m}'(m_i - m_i^*) \\ \boldsymbol{\gamma}_{\mathbf{y}}' &= \boldsymbol{\alpha}_{\mathbf{y}m}'(m_i - m_i^*) \end{aligned}$$

According to (5) technical efficiency consists of four components. The first represents a time-invariant firm-specific effect, whereas the other terms reflect the interaction of m^* with time, inputs and outputs, respectively. An interesting term in expression (5) is γ_t, since it provides information about the impact of technological change on the efficiency of production, i.e., how the unobserved farm-specific factor is suited to adjust production according to the requirements of technological change.

The values of m_i^* can be simulated by (ALVAREZ et al. 2004):

$$(6)\qquad \hat{E}\left[m_i^* \mid \mathbf{y}_i^k, \mathbf{Y}_i^{-k}, \mathbf{X}_i, \boldsymbol{\delta}\right] = \frac{\frac{1}{R}\sum_{r=1}^{R} m_{i,r}^* \hat{f}\left(\mathbf{y}_i^k \mid t, m_{i,r}^*, \mathbf{Y}_i^{-k}, \mathbf{X}_i, \boldsymbol{\delta}\right)}{\frac{1}{R}\sum_{r=1}^{R} \hat{f}\left(\mathbf{y}_i^k \mid t, m_{i,r}^*, \mathbf{Y}_i^{-k}, \mathbf{X}_i, \boldsymbol{\delta}\right)},$$

where $m_{i,r}^*$ is a draw from the population of m_i^*, R is the number of draw, and $\hat{f}$ denotes the portion of the likelihood function for firm i, evaluated at the parameter estimates and the current value of $m_{i,r}^*$. The vector $\boldsymbol{\delta}$ represent all parameters to be estimated. Using capital letter for inputs and outputs indicate that the likelihood function is evaluated for all observations of firm i.

Given the estimated level of m_i^* efficiency scores can be computed by (JONDROW et al. 1982 and ALVAREZ et al. 2004):

$$(7)\qquad -\ln TE_{ij} = E\left[u_{it} \mid \varepsilon_{it}, m_i^*\right] = \frac{\sigma\lambda}{(1+\lambda)^2}\left[\frac{\phi\left(-\lambda\frac{\varepsilon_{it} \mid m_i^*}{\sigma}\right)}{\Phi\left(-\lambda\frac{\varepsilon_{it} \mid m_i^*}{\sigma}\right)} - \lambda\frac{\varepsilon_{it} \mid m_i^*}{\sigma}\right]$$

with $\lambda = \sigma_u / \sigma_v$, $\sigma^2 = \sigma_u^2 + \sigma_v^2$ and $\varepsilon_{it} = v_{it} + \ln T_{it}$.

3 EMPIRICAL IMPLEMENTATION AND ESTIMATION RESULTS

We utilized a balanced data set consisting of eight years of observations, from 1994 to 2001, on 430 Polish agricultural farms; the total number of observations was 3,440. The respective accountancy information was provided by the Polish Institute of Agricultural and Food Economics - National Research Institute (IERiGZ-PIB). We distinguished between two outputs (crop and animal production) and four inputs (land, labour, capital and intermediate inputs). Output figures represent gross crop and animal productions. These indicators are more comprehensive measures of output than sales, since they include sales, home consumption and stock changes. Since the individual figures for crop and animal production were in current values, the variables were deflated by the corresponding price indices provided by the Statistical Office in Poland (GUS var. issues, a, b).

Land input was approximated by the sum of arable land and grassland in use. Unused land was excluded in order to have a more accurate indicator of land used in production. Labour was measured by the hours of work allocated to agriculture by family and hired labour. As an indicator of capital input the total amount of farm assets (buildings, machinery, equipment) was chosen. Since the aggregate was delivered in current values we deflated the values by the price index of agricultural investment. However, even if this gives a comprehensive indicator of total capital input it is not necessarily connected to the services provided in each year. Thus, in addition we make the simplifying assumption that capital service flows are proportional to the capital stock for each farm and in each year. Intermediate inputs were approximated by total variable costs minus depreciation. The correction was conducted in order to avoid double counting. Depreciation is an imputed measure for capital which was already accounted for with the variable total farm assets. Again, since the data set contains only current cost values we deflated the series by the price index of purchased goods and services in agriculture. The definition of variables, including some descriptive statistics are provided in Table 1.

Table 1: Variable definitions and descriptive statistics

Variable	Description	Symbol	Mean	Standard deviation	Minimum	Maximum
Crop production	gross crop production, deflated	O	127.38	149.19	1.72	2384.79
Animal production	gross animal production, deflated	Y	170.12	175.27	0.02	2895.60
Labour	total hours of work allocated to agriculture by family members and hired labour	A	3823.20	1734.06	247.00	16790.00
Land	sum of arable land and grassland in use	L	15.93	15.19	1.17	191.26
Capital	total farm assets (buildings, machinery, equipment), deflated by price index of agricultural investment	K	928.71	589.41	34.13	5181.82
Intermediate inputs	total variable costs minus depreciation, deflated by price index of purchased goods and services in agriculture	V	154.30	136.20	8.97	1748.67

Source: Own estimates.

For estimation, all variables were divided by their geometric mean. Moreover, the homogeneity restriction was imposed with regard to crop production. We conducted several estimations of (4) with various assumptions regarding the error components and *m*. First, we estimated without the aggregator function *m*. This provides a pooled estimation without accounting for the panel structure of the data (model A). The panel data structure was considered in the next two estimations, which are the random effect model (model B) and the fixed effect model (C). The random effect model results from (4) by assuming that the efficiency term u_{it} varies only over firms but not over time. Additionally, it neglects the possible impact of *m*. The fixed effect estimator results from (4) by considering the impact of m_i on the constant only. The fourth approach (D) is the model developed in (4). The last estimation is an extension insofar as it accounts for possible correlation between the unobservable component (m_i*) and the level of inputs and outputs. In order to avoid this problem ALVAREZ et al. (2004) proposed to proceed like in CHAMBERLAIN (1984) and specify m_i* as a function of inputs:

(8) $$m_i^* = \tau_t \bar{t} + \boldsymbol{\tau}_x' \overline{\ln \mathbf{x}_i} + \boldsymbol{\tau}_y' \overline{\ln \mathbf{y}_i^{-k}} + \omega_i,$$

where a bar indicates group means of the variables and $\omega \sim N(0,1)$.

Instead providing a detailed discussion we will outline some general indicators which assist in choosing the most suitable approach (Table 2).

Table 2: Overall statistical indicators

	Pooled	Random effect	Fixed effect	RPM	RPM with means
Model #	A	B	C	D	E
Assumptions in (6)	$m_i^* = 0$	$m_i^* = 0$, $u_{it} = u_i$	$a_m \neq 0$, $a_{mk} = 0$, k=m, t, y, a, l, k, v	none	D with (10)
LogL	1114.25	1809.62	1690.32	1914.49	2023.63
# of parameters	30	30	459	38	44
Variance and asymmetry parameter					
σ	0.2203***	0.2763***	0.3258***	0.1553***	0.1560***
λ	1.2059***	2.2671***	2.4165***	1.3639***	1.4467***
σ_v	0.1407	0.1219	0.1246	0.0908	0.0886
σ_u	0.1696	0.2763	0.3011	0.1256	0.1275

Note: *** denote significance at $\alpha = 0.01$.
Source: Own estimates.

Since all estimates of σ and λ are significant, Table 2 provides evidence that technical inefficiency is an important aspect in Polish agriculture. However, since all estimated models yield reasonable and comparable results regarding overall statistical indicators, a selection regarding the best representation of the production possibilities is not possible at this stage. However, as the Log Likelihood of models (D) and (E) are the highest, these models appear to be the most suitable representation of the production technology. Thus, detailed information about the parameter estimates will be provided only for these two approaches (Table 3).

First, both models suggest that technical change is a relevant phenomenon in Polish agriculture. However, the estimates reveal that the initial surveyed years were characterized by technical regression ($\alpha_T < 0$), while positive effects of innovations occurred in recent years only ($\alpha_{TT} > 0$). Moreover, crop production benefited more from technical change than animal production ($\alpha_{YT} < 0$). In addition, we estimated factor using (efficiency enhancing) technological change similar in size for all inputs. Theoretical consistency requires, inter alia, that the distance function be convex in all outputs and quasi-convex in all inputs. Although, we did not test the corresponding conditions directly, we checked whether the second order derivatives of outputs and inputs have the correct signs, i. e., $\alpha_{hh} + \alpha_h^2 - \alpha_h \geq 0$, for h = Y, A, L, K, V. The conducted calculations reveal that the condition is fulfilled for all inputs and outputs. Additionally, the estimates for the means of the random parameter estimates show that the monotonicity requirements are met. The estimated distance function is non-decreasing in outputs ($\alpha_Y \geq 0$) and non-increasing in inputs ($\alpha_h \leq 0$, for h = A, L, K, V).

Moreover, the means of the random parameter estimates are consistent with empirical observations. Animal production contributed slightly more to total agricultural output than crop production. Variable costs accounted for about 60% of total production costs. Summarising the values of α_h, with h = A, L, K, V, states that the scale elasticity is approximately -1.09, i.e., indicating slightly increasing economies of scale. Moreover, the value is comparable to other analysis of Polish agricultural production (LATRUFFE, et al. 2005).

Table 3: Parameter estimates for the random coefficient model with unobservable input

	RPM	RPM with means	RPM	RPM with means	
	(D)	(E)	(D)	(E)	
Random parameter estimates			Second order effects		
Means for random parameters					
α_0	-0.1394***	-0.1540***	0.0019**	0.0029***	α_{TT}
α_T	-0.0241***	-0.0239***	-0.0074***	-0.0058***	α_{YT}
α_Y	0.5325***	0.5239***	0.0926***	0.0928***	α_{YY}
α_A	-0.1604***	-0.1894***	-0.0071***	-0.0079***	α_{AT}
α_L	-0.1932***	-0.2492***	-0.0080***	-0.0113***	α_{LT}
α_K	-0.0763***	-0.0829***	-0.0034	-0.0020	α_{KT}
α_V	-0.6586***	-0.5582***	0.0084***	0.0117***	α_{VT}
Coefficients of unobservable factor			-0.0946***	-0.0818***	α_{AA}
α_{0M}	0.1736***	0.1306***	0.0110	0.0037	α_{LL}
α_{MM}	0.0336***	0.0135***	-0.0232	0.0099	α_{KK}
α_{TM}	0.0091***	0.0063***	0.0014	-0.0155	α_{VV}
α_{YM}	-0.0360***	-0.0224***	0.1007***	0.0812***	α_{AL}
α_{AM}	-0.0268***	-0.0234***	-0.0718***	-0.0703***	α_{AK}
α_{LM}	-0.0324***	-0.0103*	0.0600***	0.0680***	α_{AV}
α_{KM}	0.0305***	0.0169***	0.0083	-0.0184	α_{LK}
α_{VM}	0.0293***	0.0154	-0.0826***	-0.0462**	α_{LV}
Mean coefficients			0.0324***	0.0345**	α_{KV}
τ_{T_bar}		-0.0926	0.0480***	0.0515***	α_{YA}
τ_{Y_bar}		0.1844***	-0.0017	-0.0250***	α_{YL}
τ_{A_bar}		0.6841***	0.0151**	0.0140**	α_{YK}
τ_{L_bar}		1.7102***	-0.0358***	-0.0316***	α_{YV}
τ_{K_bar}		0.3445***			
τ_{V_bar}		-2.8563***			

Note: *, **,*** denote significance at a =0.1, .05 and 0.01 level, respectively. No. of observations: 3,440.
Source: Own estimates.

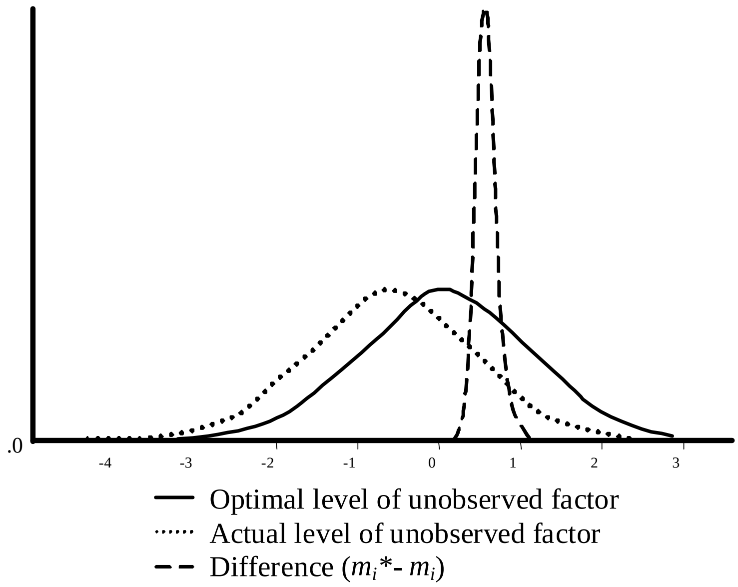
The coefficient estimates of the unobservable factor m_i* have the same structure in both approaches. Moreover, the estimated coefficients are also rather similar. Consistent with theory, both models state that the higher the factor is, the higher is the output, i.e., technical efficiency ($\alpha_{0M} > 0$, $\alpha_{MM} > 0$). The results indicate that technological change has improved productivity of the unobserved factor ($\alpha_{TM} > 0$). In addition, the unobserved component leads to an increase of production elasticities and partial factor productivities of land and labor ($\alpha_{AM} < 0$, $\alpha_{LM} < 0$,), while it has a negative impact on capital and intermediate inputs.

Considering the possibility of a correlation between the observed and unobserved inputs does not result in structurally different parameter estimates. The parameter estimates of $\boldsymbol{\tau}$ are highly significant and suggest that the unobserved component is positively correlated with farm size: m_i* becomes higher as the input of land, labor and capital increases. Only variable costs have a negative impact on the unobserved component. Moreover, since m_i* is an artificial variable, without a direct impact on input levels, the possible correlation of observable and unobservable inputs can be regarded as a minor problem (ALVAREZ et al. 2003). This interpretation is supported by the almost perfect correlation of the m_i* estimates form models (D) and (E). Thus, the following analysis will rely on the results of model (D).

4 EXPLANATION OF THE UNOBSERVED FIXED INPUT

We start the second part of our analysis by presenting some descriptive statistics with regard to the unobserved farm-specific input. We assumed in our estimation that m_i* follows a standard normal distribution. Not surprisingly, this distribution is revealed by a kernel density estimate for the factor (Figure 1). Additionally, for each farm we computed the actual level of the unobserved input, m_i, by solving (5). As Figure 1 shows, the shape of the density functions of both actual and optimal unobserved factors is the same. However, the first is shifted to the right, as expected.

Figure 1: Kernel density estimates of actual and optimal level of unobserved factor



Source: Own estimates.

4.1 Theoretical consideration

The unobserved component captures various effects on agricultural production not appropriately considered in the input-output bundle used in the estimation. These include measurement and specification errors, such as an incomplete coverage of inputs and outputs, inconsistent aggregation of farm inputs due to lack of weak separability, and unmeasured heterogeneity of the farms. Farm heterogeneity may be a result of differences in the quality of production factors, such as capital vintages, human capital, and land quality. Such systematic patterns influence farm technology, and hence cause systematic differences in long-run paths of development across the farms. In addition, m^* may be affected by determinants that are due to the organisation of agricultural production.

In the following a more systematic discussion of possible influences on m_i, m_i^* and m_i^*-m_i is conducted, in which we differentiate between scale, quality, monitoring, and diversification effects. The positive correlation of farm size on m^* obtained by model (E) suggests that farm size may have a significant impact on m^*. We capture this effect by the farms total agricultural production, averaged over the investigated period. Since the original amounts of inputs were not quality adjusted, it can be expected that quality differences will have a significant impact on the unobserved component. Our data set provides some qualitative information for land and labour, only. Regarding the first, an index of soil quality has been used. Furthermore, we assume that human capital input decreases with the age of the farmer. Younger farmers have, in general, a higher education that older ones. Our assumption neglects the impact of experience on agricultural productivity (BARTELS 1999). Indeed, given the drastic changes in the economic and institutional environment during the transition, it can be expected that formal education has become more relevant for efficient agricultural production rather than having a long practical experience.

Table 4: Definition and descriptive statistics of variables used to explain unobservable farm-specific inputs obtained by model (D)

Variable		Description	Mean	Standard deviation	Minimum	Maximum
Scale effect		Average agricultural gross output, deflated	297.51	242.98	38.48	1560.84
Factor quality	land	Index of soil quality	0.85	0.29	0.27	1.72
	labour	Average age of the household head	45.51	9.56	23.50	75.50
Farm organisation	Inputs monitoring	Share of intermediate inputs on agricultural gross output	0.54	0.08	0.32	0.97
	Labour monitoring	Share of hired labour hours on total agricultural labour input	0.04	0.06	0.00	0.55
	Land monitoring	Number of plots	5.33	4.08	1.00	42.25
Inter-sectoral diversification		Share of non-agric labour hours on total family labour	0.42	0.14	0.15	0.87
Intra-sectoral divers.	Divers. of agric. prod.	Berry-Index, based on 28 typical agricultural products	0.78	0.09	0.07	0.90
	Production intensity	Share of milk sales on total agricultural sales	0.19	0.14	0.00	0.68

Note: All variables represent average farm specific values in the investigated period (1994-2001). Number of observations: 430.

Source: Own estimates.

Polish agriculture is mainly organised in family farms. However, although family labour dominates, several farms employ a considerable amount of non-family hired labour. POLLAK (1985) and SCHMITT (1989) argue that the reasons for the dominance of family farms in Western agriculture are the transaction costs associated with the management of hired labour. The reasons for high transaction costs of hired labour result from natural uncertainties and biological production processes, both of which prevent conclusion of (almost) perfect or incentive-compatible contracts. In turn, this implies high monitoring and control costs of hired labour. With regard to family labour, these costs are expected to be much lower because of their embeddedness in agricultural households. Other monitoring efforts are associated with governing land and intermediate inputs. First, it can be presumed that fragmented farm land requires more management input and set-up times than larger plot. We could utilise information on farm-specific number of plots to control for this assumption. Second, material inputs are often regarded as substitute to labour input in conducting good agricultural practices. Moreover, this view is supported by the estimate of τ_{V_bar} reported in Table 3.

In addition, we controlled for the role of farm specialisation. Diversification of agricultural production was measured by the Berry index.[1] We assume that the more production lines have to be co-ordinated on a farm, the higher are the resources allocated to the organisation of these activities. The main reason for the higher input is the renunciation of economies of scale in management. Besides the Berry index, we also include an indicator, which is supposed to capture the effects of farm specialisation on management-intensive production activities. ALLEN and LUECK (2003) show that depending of seasonality, frequency of harvest, natural conditions and timeliness, the intensity of managerial inputs differs among the various agricultural products. They argue that especially dairy production requires intensively monitoring: a reason why milk production was less subject to industrialisation activities like those observed in poultry and hog production. In order to capture this specialisation effect we included the share of milk sales in total agricultural sales as an additional explanatory variable. Table 4 provides a summary of the independent variables as well as some descriptive statistics: The figures suggest that there is a wide variation in the socio-economic characteristics of the investigated farms, this can partly explain the unmeasured heterogeneity in the data. Moreover, since the farm business and the farm household are hardly 'separable', many factors can interact in a complex manner not necessarily fully explained by the theoretical literature. The next step of our analysis is to learn more about where the differences in the unobserved component come from, and to understand their relation to socio-economic farm-specific factors.

4.2 Empirical results

The results of the OLS estimations for m_i, m_i* and m_i*-m_i are provided in Table 5. Surprisingly, the variables discussed in Section 4.1 possess almost no explanatory power when m_i* is the regressand. The R^2 is very low, and almost no significant coefficients were obtained. Only the hypothesis regarding the diversification of agricultural production could be confirmed at the conventional level of significance. The parameter estimates for m_i are more satisfactory. The scale effect is positive, and the quality effects also have the expected signs. The same holds for inter-sectoral diversification. However, the estimates with respect to intra-sectoral diversification and farm organization are ambiguous. Diversification of production has the correct sign, however, the estimates are not significant. The opposite holds for the intensity of dairy production. The coefficients for land and labor monitoring are, contrary to our expectations, negative. However, the significance of the parameters is rather poor. Only the estimates for input monitoring, i.e., the share of material inputs in total inputs, has the correct sign and is highly significant.

Corresponding to (5), the difference of the optimal and actual value of the fixed input can be regarded as an indicator of the firm-specific effect on inefficiency. Almost all parameter estimates have the expected sign, although not all of them are significant. Inefficiency decreases with higher factor quality, and, surprisingly, with farm size. However, the effect is rather small and almost negligible. This is consistent with the findings of the random

---

[1] The index has the form BI = 1 - $\Sigma(s_{ij})^2$, where s_{ij} is the share of the j-th agricultural product in the total sales of the i-th farm.

coefficient model estimations. However, this also provides the answer to question one, raised in the introduction: The scale elasticity is approximately 1.09, which implies that rather constant economies of scale are present in the investigated sample. Thus, every farm size might be optimal, which in turn implies that scale inefficiencies should not be a severe problem in Polish agriculture, despite the dominance of rather small farms. Consistent with expectations, the parameter estimates for land and labor monitoring, despite their insignificance, suggest inefficiency increases with a higher share of hired labor and an increasing fragmentation of land. Inefficiency also increases with higher material input intensity. This might indicate that material inputs are only an insufficient substitute for other means of organizational optimization such as risk management. Because of the time constraint of agricultural households, the positive and significant estimate of inter-sectoral diversification is consistent with the theoretical considerations. The same conclusions hold for the variables that approximate farm specialization. The explanatory power in the last regression is rather low, suggesting that important aspects affecting inefficiency are not appropriately captured. However, the estimates still provide important insights about the determinants of unobserved components, i.e., firm-specific sources of inefficiency, and thus contribute to answering question 2 in the introduction regarding the factors, which drive farm efficiency.

Table 5: OLS-estimates for the unobservable farm-specific inputs obtained by model (D)

Determinants		m_i^*	m_i	$m_i^* - m_i$
Constant		-1.034*	0.199	-1.232*
Scale effect		0.000	0.002***	-0.001***
Factor quality	Land	-0.054	0.313***	-0.367**
	Labour	0.006	-0.009***	0.015***
Farm organisation	Inputs monitoring	0.022	-2.054***	2.077***
	Labour monitoring	-0.144	-0.792	0.648
	Land monitoring	0.001	-0.013*	0.014
Inter-sectoral diversification		-0.114	-1.346***	1.232***
Intra-sectoral diversification	Divers. of agric. prod.	0.870**	0.153	0.717
	Production intensity	0.288	-1.229***	1.518***
R^2		0.03	0.51	0.27
F-statistic		1.18 [10,420]	49.12*** [10,420]	17.24*** [10,420]

Note: ***, **, * indicate that the variable is significant at the 1, 5 or 10 percent level, respectively.
Source: Own estimates.

5 CONCLUSIONS

In this paper we applied the approach of Alvarez et al., (2003, 2004) for taking account of farm heterogeneity while exploring the farms' (in)efficiency. The approach utilizes a translog function and treats an unobserved farm-specific component as a random variable. The resulting econometric model is estimated as a stochastic production frontier with random coefficients (RPM). We extended the basic approach insofar as we explored the differences in the unobserved component.

The applied approach provides new insights into efficiency analysis in general, and efficiency problems faced by the Polish farms in particular. Our analysis contains at least three important implications:

First, as expected, the unobserved component model provides lower efficiency scores than the alternative approaches, such as the random or the fixed-effect model. Since the statistical properties of the RPM favor this model, our assertion that standard SFA overestimates efficiency is confirmed. At the same time, the results indicate the existence of a fifth significant, unobservable production factor besides land, capital, labor and intermediate inputs. Alvarez et al., (2004) consider this input to be managerial ability, which influences technical efficiency directly (as a farm-specific input) and indirectly (as a function) since it influences the use of other observable inputs.

Second, the empirical findings reveal that scale inefficiencies are not a severe problem in Polish agriculture. This suggests that the farms enjoy their own advantages, irrespective of their size. Thus, small farms might benefit from their flexibility, i.e., their ability to respond quickly to the dynamic changing environment (dynamic efficiency), whereas relatively large farms are likely to benefit from economies of scale in purchasing, producing and marketing operations, as well as from positive effects from innovations (static efficiency).

Third, when analyzing the differences in the unobserved component, some inefficiency sources could be identified. Since Alvarez et al. (2004), consider m_i* as optimal management (fixed level of management defining the farm's frontier), we regressed the estimates of m_i* against several variables which are, theoretically, supposed to be related to managerial skills. However, we do not find noteworthy statistical support for their conjecture. One reason might be the weak separability between the farm business and the farm household; many factors can interact in a complex and interdependent manner not fully captured by our rather simplified estimation. Thus, our estimates may be biased and the true relationship would only be revealed using an approach that explicitly takes into account the different links between the variables. On the other hand, results regarding the actual input of the unobserved component m_i provided expected and reliable results and confirm that the unobserved component might partially pick up the managerial issues. Nevertheless, the significant level of variables such as quality of the inputs (farm holders' age and soil quality) suggests that the unobserved component absorbs other farm-specific and time invariant factors, and hence should be considered more generally as a farm-specific level parameter.

Farm-specific technical efficiency is based upon deviations of actual from optimal management. Thus, if m_i equals m_i*, a farm is perfectly efficient. Drawing upon our results, a significant part of the farm-specific inefficiencies may be explained by systematic risk such as differences in quality of production factors. Furthermore, the positive influence of some monitoring and diversification effects suggests that the optimal (efficient) production level is harder to reach the higher is the managerial effort (amount) to govern the agribusiness (i.e., inputs or supervision-intensive production) and the more the managerial recourses are distributed to various economic activities. This suggests that specialization in agricultural production might bring some efficiency gains to the Polish farms. Another conclusion is that greater integration in factor markets (i.e., intermediate input) requires additional managerial efforts (amounts), which might be partly substituted by a higher quality of the entrepreneurship (i.e., education). Since the complexity of agribusiness operations increases with the increasing integration of the farm in factor and product markets, it is likely that managerial skills (quality) will increasingly gain in importance.

REFERENCES

ALLEN, D. W. & LUECK, D. (2003): The Nature of the Farm: Contracts, Risk, and Organization in Agriculture. MIT Press.

ÁLVAREZ A. & ARIAS C. & 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.

ÁLVAREZ A. & ARIAS C. & GREENE W., (2004): Accounting for unobservables in production models: management and inefficiency. Economic Working Papers at Centro de Estudios Andaluces E2004/72, Centro de Estudios Andaluces.

BAKUCS L.Z., LATRUFFE L., FERTŐ I.,& FOGARASI J. (2006): Technical Efficiency of Hungarian farms before and after accession Paper prepared for presentation at "Transition in Agriculture – Agricultural Economics in Transition III" conference Budapest, 2006 November 10-11.

CHAMBERLAIN G. (1984): Panel Data, in Z. Grilichas and M. Intriligator (eds.), Handbook of Econometrics, Amsterdam, North Holland.

BARTELS H. (1999): Die Struktur und die Bestimmungsgründe der Humankapitalbildung in der Landwirtschaft. Kiel: Wissenschaftsverlag Vauk.

BRÜMMER B., GLAUBEN T., THIJSSEN G., (2002). Decompositiion of Productivity Growth using Distance Function: The Case of Dairy Farms in Three European Countries, American Journal of Agricultural Economics.84 (3), pp. 628-644.

GUS (var. issues a): Statistical Yearbook of the Republic of Poland. Central Statistical Office of Poland, Warszawa.

GUS (var. issues b): Statistical Yearbook of the Polish Agriculture. Central Statistical Office of Poland, Warszawa.

JONDROW, J., C. A. K. LOVELL, I. MATEROV & P. SCHMIDT (1982): On the Estimation of Technical Inefficiency in the Stochastic Frontier Production Function Model, Journal of Econometrics 19 (2/3), 233-238.

LATRUFFE L., BALCOMBE. K., DAVIDOVA S., ZAWALINSKA K. (2004): Determinants of technical efficiency of crop and livestock farms in Poland, Applied Economics, Vol. 36, 1255–1263.

LATRUFFE, L., BALCOMBE K., DAVIDOVA S., ZAWALINSKA K. (2005): Technical and scale efficiency of crop and livestock farms in Poland: does specialization matter? Agricultural Economics 32, 281-296.

MUNROE D. (2001): Economic Efficiency in Polish Peasant Farming: An International Perspective. Regional Studies, Vol. 35.5, pp. 461- 471.

POLLAK, R. A. (1985): A Transaction Cost Approach to Families and Households, Journal of Economic Literature, Vol. 23, No. 2., pp. 581-608.

SCHMITT, G. (1989): Farms, farm households, and productivity of resource use in agriculture. European Review of Agricultural Economics 16: 257-284.