European Journal of Operational Research 285 (2020) 1174–1188

Contents lists available at ScienceDirect

European Journal of Operational Research

journal homepage: www.elsevier.com/locate/ejor

Interfaces with Other Disciplines

Measurement of technical inefficiency and total factor productivity growth: A semiparametric stochastic input distance frontier approach and the case of Lithuanian dairy farms

Tomas Baležentis[a], Kai Sun[b,*]

[a] Lithuanian Institute of Agrarian Economics, A. Vivulskio Str. 4A-13, 03220 Vilnius, Lithuania
[b] School of Economics, Shanghai University, 99 Shangda Rd., Shanghai 200444, China

ARTICLE INFO

Article history:
Received 25 April 2019
Accepted 19 February 2020
Available online 26 February 2020

Keywords:
Productivity and competitiveness
Stochastic frontier analysis
Semiparametric smooth coefficient model
Dairy farms
Lithuania

ABSTRACT

This paper presents a four-component stochastic frontier model in which the frontier function is represented by an unknown smooth input distance function, and inefficiency is decomposed into persistent and transient inefficiencies. Furthermore, the pre-truncation mean and variance of the transient inefficiency are functions of the environmental variables. By differentiating the four-component input distance frontier with respect to the time trend, total factor productivity (TFP) growth is estimated under the semiparametric smooth coefficient framework, and is decomposed into six components, i.e., technical change, scale component, allocative component, external component, efficiency change, and residual component. The empirical example focuses on the Lithuanian dairy sector with multiple outputs. Our results show that there are some persistent and transient inefficiencies in Lithuanian dairy farms. However, these farms maintained TFP growth of 2% per annum on average during 2004–2016, and much of it is attributed to the technical change and scale components.

1. Introduction

Following the seminal work of Aigner and Lovell (1977) and Meeusen and van Den Broeck (1977), who first proposed the stochastic production frontier framework using cross-sectional data, the measurement of technical inefficiency has recently been extended in two general directions. The first direction is to model the technology frontier in more flexible manners. For example, Fan, Li, and Weersink (1996) and Kumbhakar, Park, Simar, and Tsionas (2007) relaxed the functional form assumption of the production frontier function, and proposed the nonparametric stochastic frontier (SF) analysis with inefficiency and noise terms. Coelli and Perelman (1999) used linear programming, corrected ordinary least squares (OLS), and data envelopment analysis (DEA) to estimate distance functions. More recently, Sun and Kumbhakar (2013) and Yao, Zhang, and Kumbhakar (2019) proposed the semiparametric smooth coefficient (SPSC) stochastic production frontier model, in which the input elasticities (i.e., coefficients of logged inputs) are unknown smooth functions of some non-traditional inputs—they can be viewed as firm characteristics, policy variables, or any variables that describe the production environment, and inefficiency is also modelled as a function of these non-traditional inputs (i.e., environmental variables; hereafter, Z variables).

The second direction is to model the error components of an SF model in different ways. Instead of estimating an SF model with two error terms, i.e., noise and inefficiency, with the help of panel data, Colombi, Kumbhakar, Martini, and Vittadini (2014), Kumbhakar, Lien, and Hardaker (2014), and Tsionas and Kumbhakar (2014) considered four-component stochastic production frontier models, in which there are four error components, i.e., noise, random firm-effects, persistent (i.e., time-invariant, or long-run) inefficiency, and transient (i.e., time-varying, or short-run) inefficiency. More recently, Badunenko and Kumbhakar (2017) proposed a four-component cost frontier model, and Lai and Kumbhakar (2018) suggested the use of maximum simulated likelihood to estimate a four-component production frontier model. All of these four component models have fully parametric specifications of the frontier function, and therefore are subject to possible model misspecification, at least in the frontier part of the technology.

To the best of our knowledge, this paper is the first to propose a four-component SF model in which the frontier function is represented by an unknown smooth input distance function (IDF) that is estimated nonparametrically. In other words, we relax the

* Corresponding author.
E-mail address: ksun1@shu.edu.cn (K. Sun).

https://doi.org/10.1016/j.ejor.2020.02.032
0377-2217/

parametric functional form assumption of the IDF to be estimated. Furthermore, both the pre-truncation mean and variance of the transient inefficiency are functions of the Z variables, while only the pre-truncation variance is a function of Z in Sun and Kumbhakar (2013) and Lai and Kumbhakar (2018); that is, the Z variables are the determinants of the transient inefficiency (Kumbhakar & Sun, 2013; Wang & Schmidt, 2002). While the inputs (outputs) could be argued as endogenous in a production (cost) function, the IDF is free from these endogeneity problems so long as we restrict the returns to scale (RTS) to be a constant; that is, firm- and time-invariant (Kumbhakar, 2013). Sun, Kumbhakar, and Tveterås (2015) proposed a cost frontier model in which inefficiency is decomposed into persistent and transient inefficiencies, and noise is separated from firm effects. However, instead of estimating the cost function nonparametrically, Sun et al. (2015) specified the output elasticity and input shares of the cost function as unknown smooth functions of time trend only, while no determinant of the transient inefficiency (i.e., the inclusion of the Z variables) was allowed.

In addition to proposing the novel four-component input distance frontier model with persistent and transient inefficiencies, this paper seamlessly proceeds with estimating and decomposing total factor productivity (TFP) growth. Different approaches and techniques have been available to measure TFP growth (Chambers, 1988; Coelli, Rao, and Battese, 1998; Diewert, 1981, among others). This paper follows the spirit of Kumbhakar and Sun (2012) and decomposes TFP growth by differentiating the four-component input distance frontier with respect to the time trend. Then, the resulting growth formulation of the original input distance frontier is used, along with the Divisia definition of TFP growth, to decompose TFP growth into the following six components: technical change (TC), scale component, allocative component, external component, efficiency change (EC), and residual component. In contrast, Kumbhakar and Sun (2012) decomposed TFP growth by assuming that all firms are fully technically efficient. The consequence is that the EC cannot be obtained because it is equal to zero by assumption, and the external component cannot be obtained because no Z variable is introduced in their model. However, with the modelling of transient inefficiency and its determinants (i.e., the Z variables), this paper derives a finer decomposition of TFP growth with two more components than that in Kumbhakar and Sun (2012), i.e., the EC and external component. Furthermore, there is only one output in the application of Kumbhakar and Sun (2012); therefore, no output price data is needed to compute the actual revenue share, which is equal to unity in the single-output case. However, there are two outputs in the application of this paper, and with input and output price data, we are able to compute the actual revenue shares and actual cost shares. With this share information, we can obtain non-zero scale and allocative components, by allowing for possible market power and input misallocation, respectively.

The empirical application focuses on farm-level analysis of Lithuanian dairy sector. Indeed, productivity of dairy sector in Europe and other regions has received much attention in the literature (Ang & Oude Lansink, 2017; Cechura, Grau, Hockmann, Levkovych, & Kroupova, 2017; Kumbhakar & Heshmati, 1995; Latruffe, Ureta, Carpentier, Desjeux, & Moreira, 2017; Mennig & Sauer, 2019; Sipiläinen, Kumbhakar, & Lien, 2013; Skevas, Emvalomatis, & Brümmer, 2017; 2018a; 2018b) due to several reasons. First, dairy sector, as well as the agricultural sector in general, falls under the regulatory and support policies in most countries (Kuipers, Malak-Rawlikowska, Stalgiene, & Klopcic, 2017; de Lauwere, Malak-Rawlikowska, Stalgiene, Klopcic, & Kuipers, 2018). Second, the globalised markets of dairy products imply transformations of the supply chains, including restructuring of farms. This issue is of particular importance in the Central and Eastern European countries where de-collectivisation took place in the early-1990s and created an unsustainable farm structure. The case of Lithuania provides an interesting example of a dairy sector in transition (Verhees, Malak-Rawlikowska, Stalgiene, Kuipers, & Klopčič, 2018) as the small farms are phased out due to increasing competition and the resulting structural shifts require further economic analysis. However, there has been no research on the analysis of TFP growth in the Lithuanian dairy sector based on farm-level data and especially using the SPSC approach.[1] This approach is important in taking account of differences in the marginal productivity of agricultural inputs resulting from non-optimal use of inputs. In the context of economic transformations, such situations may emerge due to input market distortions (e.g., Lithuania saw collectivisation and subsequent land reform).

In this paper, we use the data from the Farm Accountancy Data Network (FADN) that covers the period of 2004–2016, and consider both specialised milk farms and mixed milk-cattle farms. First, the technical inefficiency and marginal effects of its determinants are estimated via the nonparametric stochastic input distance frontier function. Second, the decomposition of TFP growth is carried out to identify its sources in Lithuanian dairy farms by estimating an SPSC form of the frontier function.

The rest of this paper is organized as follows. Section 2 derives the IDF from the transformation function, and proposes a four-component nonparametric input distance frontier model with persistent and transient inefficiencies. Section 3 derives the SPSC growth formulation of the input distance frontier model, and decomposes TFP growth using the growth formulation. The data that we used are presented in Section 4. The results are presented in Section 5. Finally, Section 6 concludes.

2. Estimation of an input distance frontier

Following Kumbhakar (2013), we derive the IDF as a representation of production technology from the transformation function:[2]

$$A \cdot F(Y, X^{*}; Z, t, \omega) = 1, \tag{1}$$

where $A > 0$ is the productivity parameter, and $Y \in \mathbb{R}_{+}^{Q}$ is a vector of the actual outputs.[3] $X^{*} \in \mathbb{R}_{+}^{K}$ is a vector of minimum feasible inputs, and $X^{*} = X/D$, where X is a K-vector of actual inputs, and $D \geq 1$ is the scalar distance by which the input vector, X, can be deflated such that it reaches X^{*}. $\ln D \geq 0$ is interpreted as the input-oriented technical inefficiency, which may have two components, i.e., persistent and transient inefficiencies, with the help of panel data. We will return to this point later. $Z \in \mathbb{R}^{P}$ is a vector of firm characteristics (e.g., firm age, size, etc.), t is the time trend, and ω denotes unobserved heterogeneity (e.g., individual effects). Z, t, and ω are considered exogenous. The IDF is then obtained by imposing the restriction of homogeneity of degree one in X^{*} on the transformation function (1), using the first optimal input, X_1^{*}, as the numeraire:

$$A \cdot F(Y, \widetilde{X}^{*}; Z, t, \omega) = 1/X_1^{*}, \tag{2}$$

where $\widetilde{X}^{*}$ is a vector of input ratios, with elements $\widetilde{X}_k^{*} = X_k^{*}/X_1^{*}$, $\forall k = 2, \ldots, K$. Using the fact that $\widetilde{X}_k^{*} = \widetilde{X}_k = X_k/X_1$, $\forall k = 2, \ldots, K$, and $X_1^{*} = X_1/D$, the IDF (2) can be rewritten as

$$A \cdot F(Y, \widetilde{X}; Z, t, \omega) = D/X_1, \tag{3}$$

[1] Baležentis, Li, and Baležentis (2015) applied a semiparametric approach to analysing the patterns of technical efficiency of Lithuanian dairy farms.

[2] For an alternative derivation of the IDF, see Shephard (1953), Shephard (1970), and Kumbhakar and Lovell (2000), among others.

[3] In this paper, we focus on the input-oriented technical inefficiency, and therefore the actual outputs equal the maximum feasible outputs.

where $\widetilde{X}$ is a vector of input ratios, with elements $\widetilde{X}_k$, $\forall k = 2, \ldots, K$. The problem of this IDF is that some elements of Y and $\widetilde{X}$ in $F(\cdot)$ might be endogenous. To guarantee that all the variables in $F(\cdot)$ are exogenous,[4] we follow Kumbhakar (2013) and further impose the constant returns to scale (CRS) restriction—the RTS equals one—on (3); that is, (3) is homogenous of degree minus one in Y, using the first output, Y_1, as the numeraire (Coelli & Perelman, 1999), and we get

$$A \cdot F(\widetilde{Y}, \widetilde{X}; Z, t, \omega) = (D/X_1) \cdot Y_1, \tag{4}$$

where $\widetilde{Y}$ is a vector of output ratios, with elements $\widetilde{Y}_q = Y_q/Y_1$, $\forall q = 2, \ldots, Q$.

Taking the natural logarithm for both sides of (4) and imposing an additive structure of individual effects, ω, gives

$$\ln Y_1 - \ln X_1 = f(\ln \widetilde{Y}, \ln \widetilde{X}; Z, t) + \omega + \ln A - \ln D, \tag{5}$$

where $\ln F(\widetilde{Y}, \widetilde{X}; Z, t, \omega) = f(\ln \widetilde{Y}, \ln \widetilde{X}; Z, t) + \omega$.[5]

To make (5) an estimable equation, and to estimate the input-oriented technical inefficiency, we rewrite (5) by adding the subscripts i and t as

$$\ln Y_{1it} - \ln X_{1it} = f(\ln \widetilde{Y}_{it}, \ln \widetilde{X}_{it}; Z_{it}, t) + \omega_i + v_{it} - (u_{it} + \eta_i), \tag{6}$$

where i and t index firm and year, respectively, $\ln A_{it} = v_{it}$ is the noise term, and $\ln D_{it} = u_{it} + \eta_i$ is the non-negative input-oriented technical inefficiency, since $\ln D_{it} \geq 0$. $\ln D_{it}$ is decomposed into two components, where $u_{it} \geq 0$ is the transient inefficiency that represents the inefficiency component that changes over time, and $\eta_i \geq 0$ is the persistent inefficiency that captures the inefficiency component that is time-invariant. ω_i are the firm effects. (6), therefore, is a four-component SF model in which the four error components are $[\omega_i + v_{it} - (u_{it} + \eta_i)]$; see Badunenko and Kumbhakar (2017) and Lai and Kumbhakar (2018) for fully parametric four-component stochastic cost and production frontier models, respectively.

More specifically, Badunenko and Kumbhakar (2017) imposed the translog functional form on the cost function, and Lai and Kumbhakar (2018) imposed the Cobb-Douglas functional form on the production function. However, in this paper, we relax the functional form assumptions and estimate our four-component stochastic input distance frontier when the functional form of $f(\cdot)$ is unknown. That is, we choose to use the nonparametric approach to estimating $f(\cdot)$ as an unknown smooth function. We propose a four-step estimation procedure to estimate (6), the four-component stochastic input distance frontier, including the persistent and transient technical inefficiencies, as follows.

Step 1: Rewrite (6) as

$$\ln Y_{1it} - \ln X_{1it} = f(\ln \widetilde{Y}_{it}, \ln \widetilde{X}_{it}; Z_{it}, t) - E(u_{it}|Z_{it}, t) - E(\eta_i) + e_{it}, \tag{7}$$

where

$$e_{it} = v_{it} - u_{it} + \omega_i - \eta_i + E(u_{it}|Z_{it}, t) + E(\eta_i), \tag{8}$$

and $E(e_{it}|X_{it}, Y_{it}, Z_{it}, t) = 0$ given that: (1) $E(v_{it}|X_{it}, Y_{it}, Z_{it}, t) = E(\omega_i|X_{it}, Y_{it}, Z_{it}, t) = 0$, (2) $E(u_{it}|X_{it}, Y_{it}, Z_{it}, t) = E(u_{it}|Z_{it}, t)$; that is, u_{it} is mean independent of X_{it} and Y_{it}, conditional on Z_{it} and t, and (3) $E(\eta_i|X_{it}, Y_{it}, Z_{it}, t) = E(\eta_i)$, $\forall i$ and t. We can then view $f(\cdot)$ and $E(u_{it}|Z_{it}, t)$ as two unknown functions, and view $E(\eta_i)$ as a constant, and estimate (7) via the series regression.[6] Obtain the residuals, $\hat{e}_{it}$.

Step 2: Rewrite (8) as

$$e_{it} = \chi_{0i} + \chi_{it}, \tag{9}$$

where $\chi_{0i} = E(\eta_i) + \omega_i - \eta_i$, and $\chi_{it} = E(u_{it}|Z_{it}, t) + v_{it} - u_{it}$. Then, replace e_{it} in (9) with $\hat{e}_{it}$ from Step 1, and estimate (9) as a fixed effects panel data model without any regressors, and obtain the fitted values of χ_{0i}, $\hat{\chi}_{0i}$, and the residuals, $\hat{\chi}_{it}$.

Step 3: Use the relationship:

$$\chi_{0i} = E(\eta_i) + \omega_i - \eta_i \tag{10}$$

to estimate the persistent inefficiency, $\eta_i \sim iidN^+(0, \sigma_\eta^2)$. In practice, we replace χ_{0i} in (10) with $\hat{\chi}_{0i}$ from Step 2. (10) can then be viewed as a standard cross-sectional stochastic frontier model with a constant term (i.e., $E(\eta_i)$) only, along with the noise term, $\omega_i \sim iidN(0, \sigma_\omega^2)$. Since $E(\eta_i) = \sqrt{2/\pi}\,\sigma_\eta$, we only need to estimate the two parameters in (10), i.e., σ_η^2 and σ_ω^2, via the maximum likelihood estimation (MLE). The persistent inefficiency can then be estimated from

$$E(\eta_i|r_{1i}) = \mu_{*i} + \sigma_* \cdot \frac{\phi(\mu_{*i}/\sigma_*)}{\Phi(\mu_{*i}/\sigma_*)}, \tag{11}$$

where $r_{1i} = \omega_i - \eta_i$ are the residuals in (10), $\mu_{*i} = -r_{1i}\sigma_\eta^2/\sigma^2$, $\sigma^2 = \sigma_\eta^2 + \sigma_\omega^2$, $\sigma_* = \sigma_\eta\sigma_\omega/\sigma$, and ϕ and Φ denote the standard normal density and distribution functions, respectively (Jondrow, Lovell, Materov, & Schmidt, 1982), and the persistent TE score can be computed from

$$TE_{0i} = E[\exp(-\eta_i)|r_{1i}] = \frac{\Phi(\mu_{*i}/\sigma_* - \sigma_*)}{\Phi(\mu_{*i}/\sigma_*)} \cdot \exp\left(-\mu_{*i} + \frac{1}{2}\sigma_*^2\right) \tag{12}$$

(Battese & Coelli, 1988).

Step 4: Use the relationship:

$$\chi_{it} = E(u_{it}|Z_{it}, t) + v_{it} - u_{it} \tag{13}$$

to estimate the transient inefficiency, $u_{it} \sim iidN^+(\mu(Z_{it}, t), \sigma_u^2(Z_{it}, t))$,[7] where $\mu(Z_{it}, t) = c_0 + \delta_Z' Z_{it} + \delta_t t$ and $\sigma_u(Z_{it}, t) = \exp(c_1 + \rho_Z' Z_{it} + \rho_t t)$. Hence,

$$E(u_{it}|Z_{it}, t) = \frac{\mu(Z_{it}, t)a(Z_{it}, t)}{2} + \frac{\sigma_u(Z_{it}, t)a(Z_{it}, t)}{\sqrt{2\pi}} \cdot \exp\left(-\frac{1}{2}\left(\frac{\mu(Z_{it}, t)}{\sigma_u(Z_{it}, t)}\right)^2\right), \tag{14}$$

where $a(Z_{it}, t) = [\Phi(\mu(Z_{it}, t)/\sigma_u(Z_{it}, t))]^{-1}$ (Kumbhakar and Lovell, 2000, Chapter 3). In practice, we replace χ_{it} in (13) with $\hat{\chi}_{it}$ from Step 2. (13) can then be viewed as a non-linear panel SF model with a non-linear frontier function $E(u_{it}|Z_{it}, t)$, along with the noise term, $v_{it} \sim iidN(0, \sigma_v^2)$.[8] We can then estimate all the parameters in (13)—that is, c_0, c_1, δ_Z, ρ_Z, δ_t, ρ_t, and σ_v^2—via the MLE. Let

[4] Following Kumbhakar (2013), the RTS is required to be a constant for all the input and output ratios to be exogenous. According to conventional wisdom, a natural choice of this constant is unity.

[5] If we rewrite (5) as

$$\ln D = \ln X_1 - \ln Y_1 + f(\ln \widetilde{Y}, \ln \widetilde{X}; Z, t) + \omega + \ln A,$$

we can then get, under the IDF framework, the input elasticities as

$$\frac{\partial \ln D}{\partial \ln X_1} = 1 - \sum_{k=2}^{K} \frac{\partial f(\cdot)}{\partial \ln \widetilde{X}_k}; \quad \frac{\partial \ln D}{\partial \ln X_k} = \frac{\partial f(\cdot)}{\partial \ln \widetilde{X}_k}, \ \forall k = 2, \ldots, K,$$

and the output elasticities as

$$\frac{\partial \ln D}{\partial \ln Y_1} = -1 - \sum_{q=2}^{Q} \frac{\partial f(\cdot)}{\partial \ln \widetilde{Y}_q}; \quad \frac{\partial \ln D}{\partial \ln Y_q} = \frac{\partial f(\cdot)}{\partial \ln \widetilde{Y}_q}, \ \forall q = 2, \ldots, Q,$$

respectively. In other words, we can link the $\ln D$ with $f(.)$.

[6] The *gam* function of the *mgcv* package in R comes in handy to do this. The codes are available from the authors upon request.

[7] Note that the truncated normal distribution nests some more parsimonious distributions, such as the exponential distribution, as a special case (Meesters, 2014).

[8] Sun et al. (2015) proposed a similar decomposition and estimation of the transient inefficiency, without allowing the transient inefficiency to be determined by the external factors, Z_{it}, and time trend.

$\mu_{it} = \mu(Z_{it}, t)$, and $\sigma^2_{uit} = \sigma^2_u(Z_{it}, t)$, the transient inefficiency is estimated from

$$E(u_{it}|r_{2it}) = \tilde{\mu}_{it} + \tilde{\sigma}_{*it} \cdot \frac{\phi(\tilde{\mu}_{it}/\tilde{\sigma}_{*it})}{\Phi(\tilde{\mu}_{it}/\tilde{\sigma}_{*it})}, \tag{15}$$

where $r_{2it} = v_{it} - u_{it}$ are the residuals in (13), $\tilde{\mu}_{it} = (\mu_{it}\sigma^2_v - r_{2it}\sigma^2_{uit})/\tilde{\sigma}^2_{it}$, $\tilde{\sigma}^2_{it} = \sigma^2_v + \sigma^2_{uit}$, and $\tilde{\sigma}_{*it} = \sigma_v\sigma_{uit}/\tilde{\sigma}_{it}$ (Jondrow et al., 1982), and the transient TE score is computed from

$$TE_{it} = E[\exp(-u_{it})|r_{2it}] = \frac{\Phi(\tilde{\mu}_{it}/\tilde{\sigma}_{*it} - \tilde{\sigma}_{*it})}{\Phi(\tilde{\mu}_{it}/\tilde{\sigma}_{*it})} \cdot \exp\left(-\tilde{\mu}_{it} + \frac{1}{2}\tilde{\sigma}^2_{*it}\right) \tag{16}$$

(Battese & Coelli, 1988). Finally, the inefficiency change,

$$\begin{aligned}\partial u_{it}/\partial t = \partial E(u_{it}|r_{2it})/\partial t = \delta_t\left[\frac{\sigma^2_v}{\tilde{\sigma}^2_{it}}(1 - m_{it}g_{it} - g^2_{it})\right] \\ + \frac{\rho_t}{\tilde{\sigma}^2_{it}}\left\{\sigma^2_v\,\tilde{\sigma}_{*it}\left[g_{it}(1 + m^2_{it}) + m_{it}g^2_{it}\right]\right. \\ \left. - 2\tilde{\sigma}^2_{*it}(r_{2it} + \mu_{it})(1 - g^2_{it} - m_{it}g_{it})\right\},\end{aligned} \tag{17}$$

where $m_{it} = \tilde{\mu}_{it}/\tilde{\sigma}_{*it}$ and $g_{it} = \phi(m_{it})/\Phi(m_{it})$, is estimated using the formula given in Kumbhakar and Sun (2013).

3. A growth formulation and TFP growth

Based on (6), we can then proceed with estimating and decomposing TFP growth as follows. We first re write (6) as

$$\eta_i + u_{it} + \ln Y_{1it} - \ln X_{1it} = f(\ln\widetilde{Y}_{it}, \ln\widetilde{X}_{it}; Z_{it}, t) + \omega_i + v_{it}. \tag{18}$$

Then follow Kumbhakar and Sun (2012) and take the time derivative of both sides of (18), and we would have

$$\begin{aligned}\frac{\partial u_{it}}{\partial t} + \dot{Y}_{1it} - \dot{X}_{1it} = \beta_0(\cdot) + \sum_{k=2}^{K}\beta_k(\cdot)\dot{\widetilde{X}}_{kit} + \sum_{q=2}^{Q}\gamma_q(\cdot)\dot{\widetilde{Y}}_{qit} \\ + \sum_{p=1}^{P}\varphi_p(\cdot)\nabla_t Z_{pit} + \nu_{it},\end{aligned} \tag{19}$$

where $\dot{Y}_{1it} = \partial\ln Y_{1it}/\partial t$, $\dot{X}_{1it} = \partial\ln X_{1it}/\partial t$, $\dot{\widetilde{X}}_{kit} = \partial\ln\widetilde{X}_{kit}/\partial t$, $\forall k = 2,\ldots,K$, $\dot{\widetilde{Y}}_{qit} = \partial\ln\widetilde{Y}_{qit}/\partial t, \forall q = 2,\ldots,Q$, and $\nabla_t Z_{pit} = \partial Z_{pit}/\partial t$, $\forall p = 1,\ldots,P$. In addition, we interpret the regression coefficients as follows:

$$\beta_0(\cdot) = \frac{\partial f(\cdot)}{\partial t}; \quad \beta_k(\cdot) = \frac{\partial f(\cdot)}{\partial\ln\widetilde{X}_{kit}}; \quad \gamma_q(\cdot) = \frac{\partial f(\cdot)}{\partial\ln\widetilde{Y}_{qit}};$$

$$\varphi_p(\cdot) = \frac{\partial f(\cdot)}{\partial Z_{pit}}.$$

The properties of the IDF indicates that $f(\cdot)$ is non-decreasing in $\ln\widetilde{X}$ (i.e., $\beta_k(\cdot)\geq 0$), and non-increasing in $\ln\widetilde{Y}$ (i.e., $\gamma_q(\cdot)\leq 0$). Technical progress means that $f(\cdot)$ is increasing in t. The linear homogeneity property of the IDF indicates that $\beta_1(\cdot) \equiv 1 - \sum_{k=2}^{K}\beta_k(\cdot)$, and the CRS restriction indicates that $\gamma_1(\cdot) = -1 - \sum_{q=2}^{Q}\gamma_q(\cdot)$.[9] These functional coefficients have clear economic meanings. We can interpret $\beta_0(\cdot)$ as technical change (TC) because it measures the shift of the input distance frontier over time, *ceteris paribus*. The rest are various elasticities, and $\nu_{it} = \partial v_{it}/\partial t$ is the mean-zero random noise. In practice, we can replace the true $\partial u_{it}/\partial t$ in (19) with its estimated counterpart from (17), and the other right-hand-side variables in (19) are essentially calculated growth rates via log differences from the data.

The growth formulation allows us to decompose TFP growth into several components. To see this, we follow Kumbhakar and Sun (2012) and start with the Divisia definition of TFP growth: $T\dot{F}P_{it} \equiv \sum_{q=1}^{Q}R_{qit}\dot{Y}_{qit} - \sum_{k=1}^{K}S_{kit}\dot{X}_{kit}$, where R_{qit} denotes the revenue share of each output ($q = 1,\ldots,Q$), and S_{kit} the cost share of each input ($k = 1,\ldots,K$), $\forall i$ and t. Add $T\dot{F}P_{it}$ to both sides of (19) and re arrange, and we can show that

$$\begin{aligned}T\dot{F}P_{it} = \beta_0(\cdot) + \sum_{q=2}^{Q}(R_{qit} + \gamma_q(\cdot))\dot{\widetilde{Y}}_{qit} + \sum_{k=2}^{K}(\beta_k(\cdot) - S_{kit})\dot{\widetilde{X}}_{kit} \\ + \sum_{p=1}^{P}\varphi_p(\cdot)\nabla_t Z_{pit} - \frac{\partial u_{it}}{\partial t} + \nu_{it},\end{aligned} \tag{20}$$

using the fact that $\sum_{q=1}^{Q}R_{qit} = 1$, $\sum_{k=1}^{K}S_{kit} = 1$, $\dot{\widetilde{Y}}_{1it} = 0$ and $\dot{\widetilde{X}}_{1it} = 0$, $\forall i$ and t. Thus, TFP growth has six components.[10] The first component is TC captured by $\beta_0(\cdot)$. The second component is the scale component: $\sum_{q=2}^{Q}(R_{qit} + \gamma_q(\cdot))\dot{\widetilde{Y}}_{qit} = \sum_{q=1}^{Q}(R_{qit} + \gamma_q(\cdot))\dot{Y}_{qit}$. If perfect competition and CRS hold at the same time, then the scale component will be zero.[11] The third component usually refers to the allocative component ($\sum_{k=2}^{K}(\beta_k(\cdot) - S_{kit})\dot{\widetilde{X}}_{kit}$) because it captures the effects of input misallocation (i.e., deviation of input bundle from the optimal). If producers minimise their costs and allocate their inputs optimally, this allocative component will be zero.[12] The fourth component is the external component ($\sum_{p=1}^{P}\varphi_p(\cdot)\nabla_t Z_{pit}$) that captures the effect of external factors such as the scope of production, farm age and size, degree of competition, and other variables that can affect production and cannot be classified as traditional inputs (e.g., capital and labour) and outputs. The fifth component ($-\partial u_{it}/\partial t$) measures the efficiency change (EC). A positive value of it indicates inefficiency diminution over time, *ceteris paribus*. Finally, the last component is the residual component that can be viewed as a productivity shock that is not explained by the model.

4. Data

Our empirical research relies upon farm-level panel data from the FADN[13] describing the performance of Lithuanian family farms. We look into the performance of these dairy farms and consider the two types of farms: specialised dairy farms and mixed farms producing both milk and beef (farming types 45 and 47 as defined by the European Commission Regulation 1242/2008).[14] The years

---

[9] In fact, $\gamma_q(\cdot) = \partial\ln D/\partial\ln Y_q$, $\forall q = 1,\ldots,Q$.

[10] Since Kumbhakar and Sun (2012) did not introduce inefficiency and Z variables in their model, they decomposed TFP growth into four components only, i.e., without the efficiency change and external component. See Feng and Serletis (2010) and Restrepo-Tobon, Kumbhakar, and Sun (2015) for alternative decompositions of TFP growth.

[11] This is because the scale component can be further decomposed as $(1 - RTS)\sum_{q=1}^{Q}\gamma_q(\cdot)\dot{Y}_q + \sum_{q=1}^{Q}\left(R_q - \frac{\gamma_q(\cdot)}{\sum_{q=1}^{Q}\gamma_q(\cdot)}\right)\dot{Y}_q$, where $RTS = -1/\sum_{q=1}^{Q}\gamma_q(\cdot)$, R_q is the *actual* revenue share, and $\frac{\gamma_q(\cdot)}{\sum_{q=1}^{Q}\gamma_q(\cdot)}$ is the *shadow* revenue share. Perfect competition implies that the *actual* revenue share equals the *shadow* revenue share. Furthermore, CRS implies that $RTS = 1$, or equivalently, $\sum_{q=1}^{Q}\gamma_q(\cdot) = -1$. Therefore, if perfect competition and CRS hold at the same time, $R_q = -\gamma_q(.)$, $\forall q = 1,\ldots,Q$, and the scale component will then become zero.

[12] Following Kumbhakar and Sun (2012), if the producers minimise their costs, $W'X$, where W is the input price vector, subject to the IDF given in (5), then we can get the first-order conditions: $\frac{W_1}{W_k} = \frac{\partial D/\partial X_1}{\partial D/\partial X_k} = \frac{\beta_1(.)/X_1}{\beta_k(.)/X_k}$, $\forall k = 2,\ldots,K$. It will then be straightforward to show that the cost share $S_k \equiv \frac{W_kX_k}{W_1X_1+\cdots+W_kX_k} = \beta_k(.)$, $\forall k = 2,\ldots,K$. Therefore, the allocative component is zero under cost minimisation.

[13] FADN is the European Union (EU)-wide system for collecting data on farm performance on an annual basis. The representative surveys are carried out within each country by applying random sampling to collect the farm-level data following unified methodology. Thus, each observation represents a certain number of farms operating in a certain country (Commission, 2019).

[14] Both of these farm types are treated as specialist holdings (animal production) in the FADN. Farms are grouped into different farm types with respect to the standard output (average value of the production based on five-year data). Annex I of

2004–2016 are covered in the analysis. Because the FADN relies on rotating sample, the resulting panel is unbalanced. After removing the outliers,[15] the sample comprises a total of 3536 farm-year observations (i.e., 1163 farms appear in the panel). In particular, 3147 observations are available as the specialised milk farms, and the other 389 observations are for mixed milk-cattle farms. The SPSC model proposed in this paper works seamlessly with unbalanced panel datasets, yet additional noise may be introduced into the empirical analysis due to the changes in the farms surveyed during the sample period.

The production technology is modelled in terms of the five input variables and two output variables, following, e.g., Cabrera, Solis, and Del Corral (2010) and Latruffe et al. (2017). As we seek to model and decompose TFP growth using (20), the input and output prices are also required. The inputs include:

1. Labour (X_1) measured in hours worked by both farmer's family members and hired labour force. The price of labour (W_1) is assumed to be equal to the price of the hired labour on a farm. In case the latter datum is missing, the annual average is applied. Therefore, the price of labour is varying across the farms in most cases;
2. Herd size (X_2) measured in livestock units (LSU).[16] The price of maintaining a single LSU (W_2) is assumed to be equal to the ratio of livestock-specific expenses to the quantity of LSU on a farm. In this setting, the price of a unit of the herd varies across the farms due to differences in veterinary expenses as well as other factors related to the differences in farming practices;
3. Intermediate consumption (X_3) – the implicit quantity index that is obtained by applying price index (W_3) provided by Eurostat (goods and services currently consumed in agriculture, base year 2010) on the value of the intermediate consumption. Intermediate consumption represents specific costs (feedingstuffs, veterinary services, and livestock insurance for livestock; and seeds, fertilizers, crop protection products, and crop insurance for crops) and overheads. In this instance, the prices are year-specific but not farm-specific;
4. Capital assets (X_4) – the implicit quantity index that is obtained by deflating the value of machinery and buildings at the beginning of the year by the price index provided by Eurostat (goods and services contributing to agricultural investment, base year 2010). Note that capital assets do not include the value of livestock to avoid double counting. The price of the capital asset (W_4) is obtained as the ratio of the interests paid and depreciation to the implicit quantity of the capital assets. Therefore, the prices of the capital assets are observation-specific; and
5. Utilised agricultural area (UAA, X_5) comprises both own and rented area measured in hectares. Land price (W_5) is obtained as the rent price for each farm. In case these data are not available, the annual average value is used instead.

The farm-specific prices are derived from the FADN sample. The use of farm-specific prices can be justified in the sense that milk price depends on the farming practices, which are related to milk quality. As regards the factor inputs, their prices might be affected by the contracts among farmers and owners/sellers of the inputs. Subject to data availability, we also use country-level input price indices for Lithuania as proxies of input prices. This setting may impact the allocative component of TFP growth.

The outputs include:

1. Milk (Y_1) measured in tons. The observation-specific prices of milk (P_1) are taken from the FADN database[17]; and
2. Other outputs (Y_2) measured as an implicit quantity index. The index is constructed by considering the other livestock output (i.e., other than milk), crop output and other output. The Tornqvist price index (P_2) is calculated by using the agricultural output price indices provided by Eurostat (base year 2010). This results in observation-specific prices.

The following three contextual variables are included in the smooth coefficients as well as the inefficiency distribution.

1. Herfindahl Hirschman Index (HHI), Z_1, measures the scope of the production (traditionally, it is used to measure the size of firms in relation to industry). In our case, HHI measures the concentration of the output-mix (or degree of farm's specialisation), instead of the market-wide measure used in the literature. HHI increases with increasing specialisation. In general, productivity is likely to increase with specialisation.[18] We obtain the HHI as the sum of squared revenue from different products produced on a certain farm.
2. The logged[19] support payments, Z_2, include both subsidies related to production activities and investments.[20] It is generally agreed that production and investment subsidies may contribute to acquisition of inputs (Minviel & De Witte, 2017). In practice, the use of the investment subsidies is more constrained than that of the production subsidies. Increasing input intensity should lead to higher productivity and efficiency. However, the efficiency of production (following investments and input use) may be dampened due to input market imperfection (increased profit margins for capital goods subject to support investments) and farmer's inability to take proper decisions on the input- and output-mix in the light of the increasing input use, as well as decreased incentives for productivity improvements. See Minviel and Latruffe (2017) and Latruffe et al. (2017) for discussion on support payments as a determinant of efficiency.
3. Farmer's age, Z_3, is included as older farmers are expected to possess higher managerial capabilities. For instance, the variable has been used as a proxy for experience in farming by Latruffe, Balcombe, Davidova, and Zawalinska (2004).

Regulation 1242/2008 provides detailed requirements of the standard output structure for each farming type. The basic requirement is that livestock standard output comprises at least 2/3 of the total standard output (with further requirements on the structure of a herd). See Commission (2019) for further details.

[15] Those observations that lie below $Q1 - 1.5 \cdot IQR$ and above $Q3 + 1.5 \cdot IQR$ are considered to be outliers and thus dropped from the sample, where $Q1$ and $Q3$ are the first and third quartiles, respectively, and IQR is the inter-quartile range of a sample. As a result, about 7.7% of the observations in the original sample were dropped.

[16] LSU is a common unit for different species and classes of livestock. For instance, one dairy cow equals 1 LSU (Commission, 2019).

[17] The dairy quota did not affect milk output in Lithuania. However, following its abolishment in 2015, milk markets saw increased volatility in milk prices, probably because milk markets in Europe became more volatile. This also affected milk market in Lithuania.

[18] However, specialisation (proxied by the HHI) may affect productivity through different channels: it can impact the input use and production technology and ensure higher productivity; but it can also increase vulnerability to shocks in the output markets and then decrease the input productivity due to underutilisation of inputs. See, e.g., de Roest, Ferrari, and Knickel (2018) for a discussion on the linkages between farm performance and specialisation.

[19] The effects of the support payments can be taken into account by introducing different variables into the econometric models. For instance, Mary (2013) used support payments in levels when analysing their impacts on TFP growth. We use the logged form for several reasons. First, this allows interpreting the results easier, i.e., the effect of growth in the support payments by 1% is obtained. Second, the use of the logged form decreases the likelihood of the occurrence of outlying observations. Third, the sign of the coefficient is interpreted in the same manner for either level or logged variable, as the support payments data are non-negative.

[20] Subsidies related to production correspond to indicator SE605 in Commission (2019). Subsidies on investments correspond to indicator SE406 in Commission (2019).

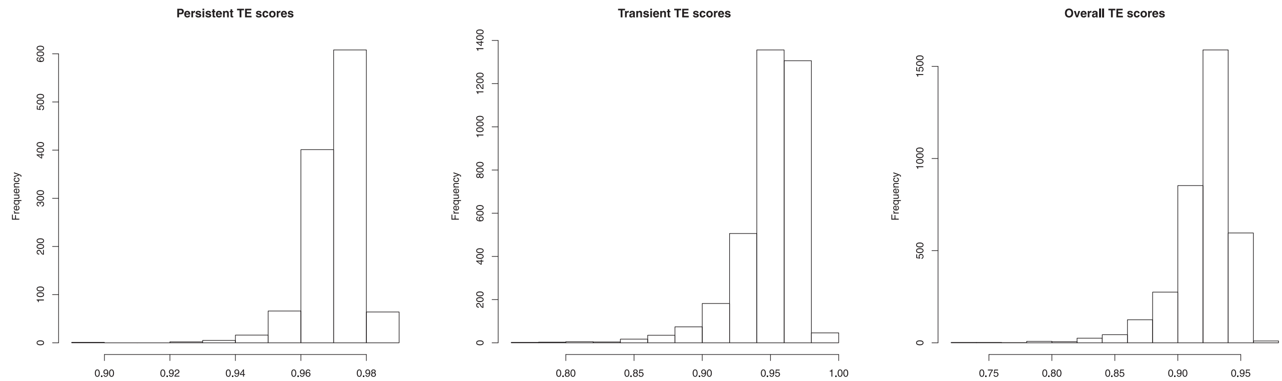

Fig. 1. TE scores.

4. Time trend t is also included to account for temporal variations in efficiency due to sector-wide shocks.

Table 1 presents the descriptive statistics, whereas Table 2 summarises the stochastic rates of growth for the input and output variables. The figures provided in Table 1 suggest that milk farms, on average, are larger than the milk-cattle farms in terms of labour input (5.9 thousand hours and 4.7 thousand hours on average, respectively). Herd size is also larger, on average, for milk farms than for milk-cattle farms (59 LSU and 48 LSU, respectively). Intermediate consumption and assets are much higher for the milk farms than for the milk-cattle farms on average. The same pattern is obtained by considering the ratios of intermediate consumption or assets to the other inputs (labour, land, or herd size). This suggests that milk farms are more likely to embark on more intensive farming practices than the milk-cattle farms. UAA is rather similar across the two farming types. Whereas the average milk output is higher for milk farms (212 tons) than for the milk-cattle farms (79 tons), there is no such difference in terms of the other outputs. Obviously, milk farms are more specialised and show higher average HHI (0.69) than the milk-cattle farms (0.51).

As suggested by Table 1, labour input shows the lowest rate of growth among all the inputs for the whole sample (1.3% per annum). The highest increase is observed for assets (8.9% per annum), which indicates that serious modernisation has been taking place in Lithuanian dairy farms. Intermediate consumption also shows rates of growth exceeding 5% per annum for both farming types. Increasing intermediate consumption may improve both quality and quantity of the outputs. As regards the quantity of outputs, both farming types show increasing production of the other outputs (e.g., meat), whereas an increase in milk production is only observed for specialised milk farms.

5. Estimation results

The application of the IDF allows us to analyse the two intertwisted issues, namely technical efficiency (TE) and TFP growth. Therefore, this section proceeds by discussing dynamics in TE and TFP of Lithuanian dairy farms. In this paper, we consider the two farming types, viz., specialised milk farms and mixed milk-cattle farms.

5.1. Technical efficiency

First, we estimate the persistent and transient TE scores as presented in (6). The mean persistent TE (PTE) score for the whole sample is 0.97 with a standard deviation of 0.01. The mean transient TE (TTE) score for the whole sample is 0.95 with a standard deviation of 0.02.[21] Fig. 1 plots the histograms of PTE and TTE,

[21] These results are quite comparable to those reported in Kumbhakar (2013), who studied the Norwegian dairy farms during 1992–2006 using the translog output dis-

Table 1
Descriptive statistics of the variables used.

	Full sample	Type = Milk	Type = Milk Cattle
	Labour (X_1), hours		
Mean	5807.98	5939.05	4747.60
SD	4960.13	5177.20	2357.41
Min	1095.00	1095.00	2030.00
Max	93544.00	93544.00	14704.00
	Herd size (X_2), LSU		
Mean	57.87	59.04	48.39
SD	69.33	71.24	50.45
Min	1.84	1.84	4.62
Max	800.77	800.77	269.46
	Intermediate consumption (X_3), index		
Mean	49673.68	51678.57	33454.23
SD	79859.02	83257.83	40028.87
Min	1752.82	1752.82	3977.43
Max	1684415.29	1684415.29	242979.70
	Assets (X_4), index		
Mean	76075.49	79606.91	47506.38
SD	121760.74	126626.43	64381.80
Min	480.35	480.35	630.63
Max	1494335.84	1494335.84	510487.32
	Utilised agricultural area (X_5), hectare		
Mean	90.39	91.65	80.14
SD	94.60	96.71	74.67
Min	2.17	2.17	4.40
Max	956.77	956.77	386.95
	Milk output (Y_1), ton		
Mean	197.64	212.28	79.16
SD	292.65	304.21	118.67
Min	1.50	3.04	1.50
Max	4867.88	4867.88	704.50
	Other outputs (Y_2), index		
Mean	37552.87	38011.36	33843.74
SD	59669.98	61651.62	40060.33
Min	98.19	98.19	3423.66
Max	1004395.41	1004395.41	262848.28
	Labour price (W_1), EUR/hour		
Mean	2.02	2.04	1.92
SD	0.66	0.67	0.58
Min	0.77	0.77	0.85
Max	6.90	6.90	5.80
	Herd price (W_2), EUR/LSU		
Mean	488.47	499.89	396.13
SD	262.93	273.45	118.13
Min	121.37	121.37	129.96
Max	11875.88	11875.88	949.98
	Intermediate consumption price (W_3), index		
Mean	1.05	1.05	1.02
SD	0.16	0.16	0.16
Min	0.77	0.77	0.77
Max	1.26	1.26	1.26
	Assets price (W_4), EUR		
Mean	0.25	0.25	0.26
SD	0.12	0.12	0.12
Min	0.01	0.01	0.06

(*continued on next page*)

Table 1 (*continued*)

	Full sample	Type = Milk	Type = Milk Cattle
Max	0.63	0.63	0.63
	Land price (W_5), EUR/hectare		
Mean	13.90	14.14	12.00
SD	12.58	12.71	11.30
Min	0.03	0.03	0.15
Max	126.41	126.41	62.52
	Milk price (P_1), EUR/ton		
Mean	230.46	233.62	204.89
SD	57.03	56.66	53.51
Min	86.89	86.89	109.73
Max	520.00	520.00	367.82
	Other outputs' price (P_2), index		
Mean	1.00	1.00	0.99
SD	0.03	0.03	0.02
Min	0.93	0.93	0.95
Max	2.41	2.41	1.06
	HHI (Z_1)		
Mean	0.67	0.69	0.51
SD	0.20	0.19	0.18
Min	0.17	0.19	0.17
Max	1.00	1.00	1.00
	Support payments in log (Z_2)		
Mean	9.35	9.34	9.37
SD	1.03	1.04	0.97
Min	6.34	6.34	7.26
Max	12.59	12.59	11.63
	Farmer's age (Z_3), years		
Mean	47.07	47.28	45.42
SD	10.46	10.32	11.37
Min	19.00	19.00	20.00
Max	75.00	75.00	75.00
	Time trend (t)		
Mean	7.54	7.63	6.85
SD	3.61	3.63	3.38
Min	1.00	1.00	1.00
Max	13.00	13.00	13.00

1. Total number of observations = 3536. 2. t is defined as year-2003, where year varies from 2004 to 2016.

respectively. There is a left-tail of the histogram of transient inefficiency, and about 14% of the observations have TTE scores of less than 0.93. This suggests that there is room for improvement in these observations. The overall TE (OTE) score is computed as PTE × TTE, and its histogram is reported in Fig. 1. The mean OTE is 0.92 with a standard deviation of 0.02. It seems that technical inefficiency associated with time-varying effects comprises a similar share of the overall technical inefficiency of Lithuanian dairy farms, compared with the persistent inefficiency that is associated with managerial practices and other stable conditions.

As suggested by (13), pre-truncation mean and variance, and also the standard deviation of the underlying truncated normal distributions of transient inefficiency are modelled with respect to explanatory variables, Z. Table 3 presents the estimates of parameters governing $\mu(Z, t)$ and $\sigma_u(Z, t)$. Out of the four variables, support payments appear to be significant at the 1% level, and farmer's age and time trend appear to be significant at the 10% level, in $\sigma_u(Z, t)$. In all instances, these variables significantly impact the standard deviation of the transient inefficiency; therefore, there is strong evidence that a change in one of these variables would change inefficiency, *ceteris paribus*. More specifically, the coefficients of support payments and time trend in $\sigma_u(Z, t)$ are significantly positive. This means that an increase in support payments or time trend would increase the pre-truncation variance of transient inefficiency, and this would also increase the post-truncation mean of inefficiency, and thus increase inefficiency itself, *holding all the other determinants and the pre-truncation mean constant*. The coefficient of farmer's age in $\sigma_u(Z, t)$ is significantly negative. This means that an increase in farmer's age would decrease the pre-truncation variance, and this would also decrease the post-truncation mean of inefficiency, and thus decrease inefficiency, *holding all the other determinants and the pre-truncation mean constant*.

tance function (ODF), IDF, and system models, and found that the mean TE scores of all these three cases are around 0.9.

Table 3
Determinants of the transient inefficiency.

Intercept	Z_1 HHI	Z_2 Support payments	Z_3 Farmer's age	t Time trend
$\mu(Z, t)$				
−0.0744	−0.0105	0.0055	0.0005*	0.0004
(0.0518)	(0.0206)	(0.0064)	(0.0003)	(0.0015)
$\sigma_u(Z, t)$				
−3.6166***	−0.1464	0.1232***	−0.0074*	0.0258*
(0.4557)	(0.2282)	(0.0471)	(0.0038)	(0.0141)

1. Z variables are described in Section 4. 2. Standard errors are in the parentheses. 3. *, ** and *** indicate significance at the level of 10%, 5% and 1%, respectively.

Given that the expected efficiency level depends on both mean and standard deviation, Kumbhakar and Sun (2013) proposed estimating the marginal effects of Z (i.e., determinants of inefficiency) to quantify the effects of Z on the expected inefficiency. Fig. 2 shows the resulting kernel densities of these marginal effects for Lithuanian family farms. The marginal coefficients obtained confirm the increasing specialisation (HHI) decreases the inefficiency as much of the probability mass falls below the zero effect. This finding suggests that specialisation of farming is beneficial for Lithuanian dairy farms. The increasing support payments contribute to increasing inefficiency. Note that this is a result of interaction between the pre-truncation mean ($\mu(Z, t)$) and pre-truncation standard deviation ($\sigma_u(Z, t)$) of the underlying inefficiency distribution. An increase in either $\mu(Z, t)$ or $\sigma_u(Z, t)$, *ceteris paribus*, would increase the post-truncation mean of inefficiency, and thus increase inefficiency itself.[22] This indicates the need for improvement in the effectiveness of the support measures under

[22] The model involving both farming type and regional dummies rendered significant coefficient of support payments at the 10% level in $\mu(Z, t)$ in Appendix A, and this provides further evidence that support payments increase inefficiency. In addition, Appendix B provides estimation results using two alternative measures of subsidies, i.e., support payments in level, and support payments per hectare. We do not use subsidies per LSU as a proxy for support intensity as much of the support received by Lithuanian dairy farms is related to crop production. For instance, only around 35 percent of the support payments related to production for dairy farms was related to animals in 2016 (see Table 27 on page 77 in Economics, 2017). The measure of subsidies relative to the output is not appropriate as it induces endogeneity (less efficient farms produce less output and the subsidy-to-output ratio increases). Further analysis of support-efficiency link is required by representing the

Table 2
Stochastic annual growth rates for inputs and outputs (%), 2004–2016.

Farm Types	Labour	Herd Size	Intermediate consumption	Assets	UAA	Milk output	Other outputs
Milk	1.2	3.0	5.2	8.5	2.7	3.8	8.1
Milk and Cattle	1.6	7.4	6.6	10.2	5.2	0.0	10.7
Whole Sample	1.3	3.5	5.4	8.9	2.9	4.0	8.4

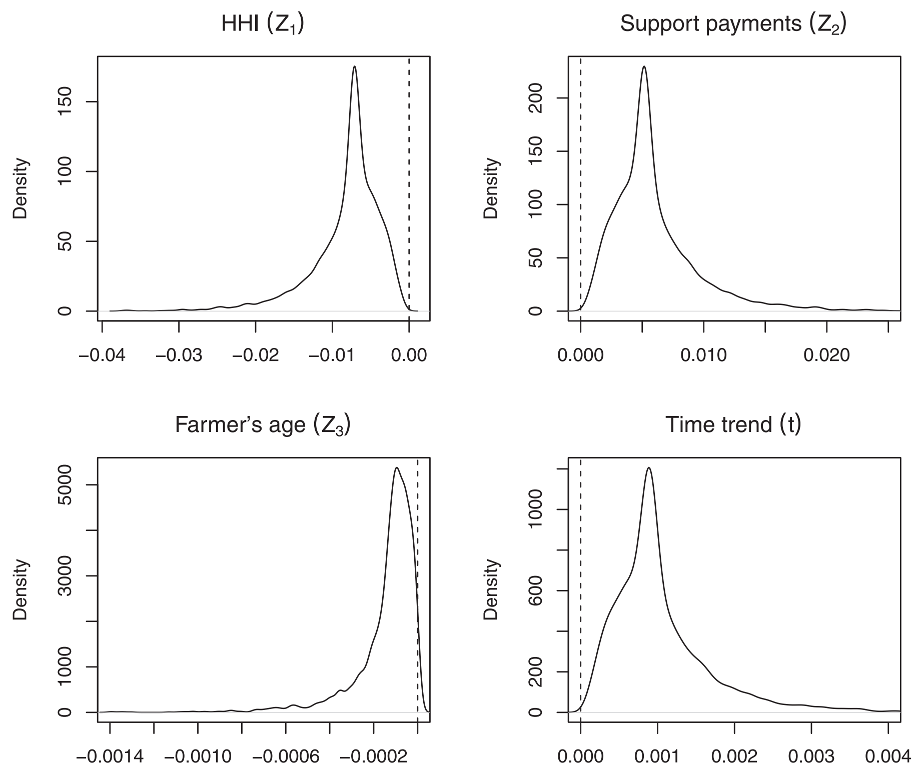

Fig. 2. Marginal effects of the determinants of transient inefficiency.

the Common Agricultural Policy (CAP). Increasing farmer's age decreases inefficiency. In the case of Lithuanian dairy farms, older farmers seem to be more experienced and applying more efficient farming practices. Finally, the inefficiency is likely to increase with time for most farms.

5.2. TFP growth

This section presents the results of estimation of (19). This allows us to look into the properties of the underlying technology by considering the regression coefficients. In addition, TFP growth can be calculated and decomposed into several components such as technical change (TC), scale, and allocative components, etc. The FADN relies on the rotating sample as some farms do not enter the survey in the next time period. Therefore, we construct the data set that only comprises observations for two consecutive time periods for each farm for estimation of TFP growth. The details reported in Section 4 imply that each farm stays in the sample for 3536/1163=3.04 years on average. Table C.1 in Appendix C presents the descriptive statistics of the variables used to estimate (19).

5.2.1. Smooth coefficients

The smooth coefficient model in (19) is estimated to proceed with the analysis of TFP growth. The regression coefficients are functions of inputs, outputs, time trend and the three contextual variables (i.e., HHI for output structure, logged support payments, and farmer's age). For ease of comparison, Fig. 3 plots the empirical cumulative distribution functions (ECDF) of the estimated smooth coefficients of milk farm type versus milk-cattle farm type. To understand these plots, we can see, for example, that the ECDF of β_0 of milk farm type lies above that of milk-cattle farm type over most of the domain. This indicates that the histogram of β_0 estimates of milk farm type is to the left of that of milk-cattle farm type, and that the β_0 (i.e., TC) estimates of milk farm type are stochastically smaller than (i.e., first-order stochastically dominated by) that of milk-cattle farm type.[23] This suggests that mixed farms enjoy higher rates of technical progress. Looking at the output structure, these changes in distributions of the TC coincides with the increasing diversification of farms (Table 2). More specifically, milk farms show a decrease in specialisation as the growth rates in the other outputs exceed those for the milk output, yet milk-cattle farms show even higher discrepancy of growth rates in the two outputs. This indirectly implies that increasing the scope of production is associated with an outward frontier shift in Lithuanian dairy farms.

The general trends indicated by the ECDFs can be represented by exact values by comparing the mean values of different coefficients across the farm types in Table 4, which presents the main distributional characteristics for the smooth coefficients.[24] For instance, the mean value of β_0 is almost twofold for the mixed farms, compared with specialised ones. However, the signs of the coefficients coincide across the farm types at means and the other quartiles.

The distribution of β_1, i.e., distance elasticity with respect to labour, is much wider for the mixed milk-cattle farms than for the specialised milk farms. Therefore, the mixed milk-cattle farms tend to show different levels of labour productivity, whereas the specialised farms show more homogeneous estimates. The distri-

levels and intensity of support in different ways in order to draw robust conclusions on the issue. This can be saved as future work.

23 To put it another way, a randomly selected milk farm type has a lower level of technical progress than a randomly selected milk-cattle farm type. However, no stochastic dominance between the two farming types is found in terms of persistent, transient, and overall TE scores, as well as marginal effects of the determinants of the transient inefficiency. The average TE scores and marginal effects of the determinants of the transient inefficiency of the two farming types are qualitatively similar.

24 Because the smooth coefficients and TFP growth components are unknown nonlinear functions of data, the distributions of these estimates might be far away from normal distributions. Therefore, it is recommended that we use the bootstrap to compute the standard errors of the mean and quartiles. For further details about bootstrap, see Cameron and Trivedi (2005, Chapter 11).

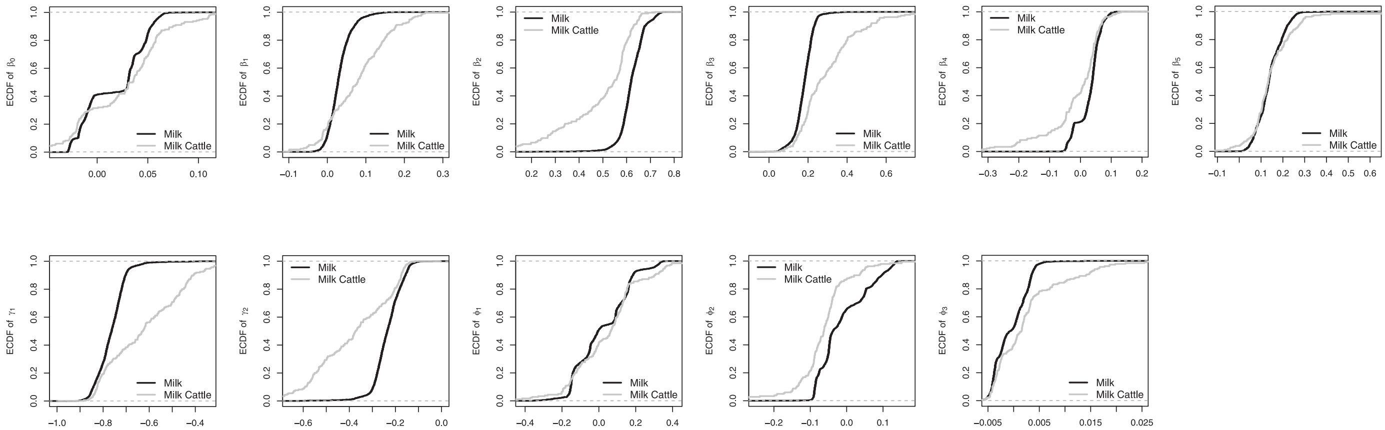

Fig. 3. ECDF plots of the smooth coefficients.

Table 4
Summary statistics of the smooth coefficients.

	β_0 TC	β_1 Labour	β_2 Herd size	β_3 Intermediate consumption	β_4 Assets	β_5 UAA	γ_1 Milk output	γ_2 Other outputs	φ_1 HHI	φ_2 Support payments	φ_3 Age
Full sample											
Mean	0.0189	0.0355	0.6080	0.1912	0.0245	0.1408	−0.7519	−0.2481	0.0218	−0.0164	−0.0002
	(0.0009)	(0.0012)	(0.0133)	(0.0048)	(0.0012)	(0.0035)	(0.0170)	(0.0058)	(0.0035)	(0.0016)	(0.0001)
Q1	−0.0116	0.0118	0.5865	0.1540	0.0087	0.0978	−0.8045	−0.2785	−0.1187	−0.0708	−0.0036
	(0.0004)	(0.0005)	(0.0008)	(0.0006)	(0.0003)	(0.0014)	(0.1093)	(0.0235)	(0.0016)	(0.0022)	(0.0000)
Q2	0.0306	0.0282	0.6143	0.1840	0.0365	0.1341	−0.7606	−0.2394	0.0003	−0.0416	−0.0002
	(0.0019)	(0.0004)	(0.0009)	(0.0009)	(0.0008)	(0.0009)	(0.0011)	(0.0012)	(0.0044)	(0.0017)	(0.0001)
Q3	0.0460	0.0504	0.6512	0.2136	0.0509	0.1813	−0.7215	−0.1955	0.1448	0.0429	0.0024
	(0.0007)	(0.0014)	(0.0815)	(0.0342)	(0.0012)	(0.0178)	(0.0007)	(0.0007)	(0.0010)	(0.0014)	(0.0000)
Type = Milk											
Mean	0.0182	0.0310	0.6210	0.1796	0.0286	0.1397	−0.7656	−0.2344	0.0198	−0.0113	−0.0005
	(0.0008)	(0.0010)	(0.0151)	(0.0042)	(0.0011)	(0.0037)	(0.0186)	(0.0057)	(0.0036)	(0.0016)	(0.0001)
Q1	−0.0114	0.0118	0.5917	0.1525	0.0144	0.0981	−0.8065	−0.2724	−0.1228	−0.0629	−0.0036
	(0.0005)	(0.0004)	(0.0009)	(0.0006)	(0.0004)	(0.0014)	(0.0967)	(0.0314)	(0.0015)	(0.0021)	(0.0000)
Q2	0.0306	0.0272	0.6179	0.1809	0.0378	0.1345	−0.7639	−0.2361	−0.0047	−0.0372	−0.0005
	(0.0021)	(0.0005)	(0.0010)	(0.0008)	(0.0008)	(0.0009)	(0.0010)	(0.0013)	(0.0043)	(0.0016)	(0.0002)
Q3	0.0452	0.0455	0.6538	0.2099	0.0515	0.1805	−0.7276	−0.1935	0.1453	0.0472	0.0022
	(0.0004)	(0.0015)	(0.0989)	(0.0314)	(0.0017)	(0.0200)	(0.0006)	(0.0007)	(0.0010)	(0.0024)	(0.0000)
Type = Milk Cattle											
Mean	0.0251	0.0786	0.4821	0.3036	−0.0152	0.1510	−0.6199	−0.3801	0.0420	−0.0653	0.0026
	(0.0040)	(0.0081)	(0.0381)	(0.0263)	(0.0070)	(0.0140)	(0.0448)	(0.0300)	(0.0154)	(0.0073)	(0.0006)
Q1	−0.0153	0.0139	0.4122	0.1936	−0.0464	0.0913	−0.7830	−0.5198	−0.1003	−0.0924	−0.0027
	(0.0011)	(0.0052)	(0.0016)	(0.0075)	(0.0025)	(0.0043)	(0.3521)	(0.1650)	(0.0054)	(0.0155)	(0.0007)
Q2	0.0350	0.0794	0.5476	0.2591	0.0154	0.1327	−0.6274	−0.3726	0.0654	−0.0623	0.0016
	(0.0039)	(0.0058)	(0.0164)	(0.0042)	(0.0055)	(0.0036)	(0.0152)	(0.0145)	(0.0159)	(0.0024)	(0.0003)
Q3	0.0547	0.1351	0.5911	0.3797	0.0432	0.2021	−0.4802	−0.2170	0.1421	−0.0359	0.0041
	(0.0054)	(0.0121)	(0.2406)	(0.1358)	(0.0073)	(0.0555)	(0.0103)	(0.0134)	(0.0102)	(0.0015)	(0.0002)

Bootstrapped standard errors are in the parentheses.

bution of β_2, i.e., distance elasticity with respect to herd size, for the specialised farms dominates that for the mixed farms. The mixed farms show higher mass of relatively low values of elasticity, as suggested by Table 4. This indicates that the specialised milk farms enjoy higher productivity of the herd than the mixed farms. Turning to β_3—elasticities for intermediate consumption—the coefficients for milk farms, once again, are much less variable than those for the milk-cattle farm type. The milk-cattle farm type exhibits wider distribution of the regression coefficients, hence distance elasticities. Furthermore, the mixed farm type first-order stochastically dominates the specialised farms in terms of β_3. In terms of β_4; that is, regression coefficients associated with asset use, the lower estimates are more likely to be observed for the mixed farms. The distributions virtually coincide for the higher values of β_4. The elasticities with respect to UAA are virtually the same across the two farming types as represented by the empirical distributions of β_5. This suggests similar levels of the utilisation of the UAA whether it is used for milk or beef cattle. The elasticity is positive across all quartiles. Thus, in general, there exists no excessive use of UAA in Lithuanian dairy farms.

The coefficients, γ's, can be interpreted as the distance elasticities with respect to outputs. Alternatively, $-\gamma_q$ shows the percent increase in the input quantity of X_1 (the numeraire input) due to a 1 percent increase in the production of Y_q, $\forall q = 1, 2$ (Kumbhakar & Sun, 2012). As suggested by Fig. 3, the milk farms have more negative values of γ_1 than milk-cattle farms. The opposite holds for γ_2. This indicates that milk farms tend to put relatively more resources than milk-cattle farms into production of one additional percent of milk output. Therefore, the quality of the outputs might vary across specialised and mixed farms due to different input intensity. This also partially explains the effect of the diversification (as represented by the marginal effects of the HHI) on the efficiency levels: specialised farms are more input-intensive and tend to generate higher efficiency. Finally, φ_p indicates the elasticities of the IDF with respect to the contextual variables, Z. These elastici-

Table 5
Summary statistics of the time derivatives of the smooth coefficients.

	$\frac{\partial\beta_0}{\partial t}$ TC	$\frac{\partial\beta_1}{\partial t}$ Labour	$\frac{\partial\beta_2}{\partial t}$ Herd size	$\frac{\partial\beta_3}{\partial t}$ Intermediate consumption	$\frac{\partial\beta_4}{\partial t}$ Assets	$\frac{\partial\beta_5}{\partial t}$ UAA	$\frac{\partial\gamma_1}{\partial t}$ Milk output	$\frac{\partial\gamma_2}{\partial t}$ Other outputs	$\frac{\partial\varphi_1}{\partial t}$ HHI	$\frac{\partial\varphi_2}{\partial t}$ Support payments	$\frac{\partial\varphi_3}{\partial t}$ Age
Full sample											
Mean	−0.0074	−0.0036	−0.0003	0.0004	0.0046	−0.0010	−0.0027	0.0027	0.0225	0.0024	−0.0002
	(0.0006)	(0.0010)	(0.0016)	(0.0014)	(0.0010)	(0.0018)	(0.0009)	(0.0009)	(0.0031)	(0.0015)	(0.0001)
Q1	−0.0307	−0.0225	−0.0388	−0.0396	−0.0304	−0.0476	−0.0344	−0.0314	−0.1024	−0.0268	−0.0024
	(0.0002)	(0.0008)	(0.0026)	(0.0013)	(0.0003)	(0.0016)	(0.0005)	(0.0006)	(0.0005)	(0.0012)	(0.0001)
Q2	−0.0033	−0.0068	−0.0015	0.0047	0.0151	0.0064	−0.0006	0.0006	0.0672	−0.0062	−0.0004
	(0.0005)	(0.0004)	(0.0012)	(0.0011)	(0.0009)	(0.0014)	(0.0012)	(0.0012)	(0.0028)	(0.0012)	(0.0000)
Q3	0.0129	0.0141	0.0258	0.0367	0.0389	0.0482	0.0314	0.0344	0.1217	0.0233	0.0015
	(0.0005)	(0.0004)	(0.0007)	(0.0015)	(0.0010)	(0.0013)	(0.0006)	(0.0005)	(0.0021)	(0.0010)	(0.0000)
Type = Milk											
Mean	−0.0066	−0.0026	0.0002	0.0002	0.0047	−0.0025	−0.0015	0.0015	0.0177	0.0034	−0.0002
	(0.0006)	(0.0008)	(0.0014)	(0.0012)	(0.0010)	(0.0015)	(0.0009)	(0.0009)	(0.0031)	(0.0015)	(0.0001)
Q1	−0.0300	−0.0217	−0.0361	−0.0386	−0.0296	−0.0490	−0.0320	−0.0316	−0.1073	−0.0256	−0.0023
	(0.0002)	(0.0008)	(0.0025)	(0.0016)	(0.0003)	(0.0017)	(0.0005)	(0.0005)	(0.0005)	(0.0012)	(0.0001)
Q2	−0.0032	−0.0066	0.0008	0.0042	0.0156	0.0047	0.0010	−0.0010	0.0646	−0.0055	−0.0004
	(0.0005)	(0.0004)	(0.0011)	(0.0012)	(0.0009)	(0.0018)	(0.0013)	(0.0013)	(0.0030)	(0.0013)	(0.0000)
Q3	0.0127	0.0129	0.0251	0.0354	0.0389	0.0474	0.0316	0.0320	0.1205	0.0232	0.0014
	(0.0006)	(0.0003)	(0.0007)	(0.0015)	(0.0010)	(0.0012)	(0.0005)	(0.0005)	(0.0021)	(0.0010)	(0.0000)
Type = Milk Cattle											
Mean	−0.0153	−0.0135	−0.0048	0.0019	0.0036	0.0128	−0.0138	0.0138	0.0686	−0.0076	−0.0002
	(0.0039)	(0.0069)	(0.0108)	(0.0078)	(0.0043)	(0.0121)	(0.0041)	(0.0040)	(0.0126)	(0.0066)	(0.0004)
Q1	−0.0398	−0.0489	−0.0638	−0.0600	−0.0386	−0.0091	−0.0474	−0.0225	0.0002	−0.0495	−0.0032
	(0.0038)	(0.0034)	(0.0069)	(0.0033)	(0.0037)	(0.0037)	(0.0029)	(0.0049)	(0.0063)	(0.0052)	(0.0003)
Q2	−0.0065	−0.0098	−0.0246	0.0240	0.0113	0.0196	−0.0108	0.0108	0.0835	−0.0145	−0.0003
	(0.0019)	(0.0036)	(0.0043)	(0.0067)	(0.0065)	(0.0025)	(0.0026)	(0.0027)	(0.0068)	(0.0034)	(0.0002)
Q3	0.0136	0.0354	0.0408	0.0656	0.0380	0.0636	0.0225	0.0474	0.1478	0.0233	0.0033
	(0.0021)	(0.0047)	(0.0049)	(0.0048)	(0.0028)	(0.0083)	(0.0049)	(0.0028)	(0.0172)	(0.0037)	(0.0005)

Bootstrapped standard errors are in the parentheses.

ties capture the effects of HHI (Z_1), log of support payments (Z_2), and farmer's age (Z_3) on the growth of the distance function. If φ_p is positive (negative), then it means that an increase in the growth of any of these variables, *ceteris paribus*, would increase (decrease) the rate of change of the distance from the producer to the frontier.

In general, the dispersion of regression coefficients in the mixed farms can be related to different productivity levels associated with different production processes and products produced there, in contrast to single product and production technology prevailing in the specialised milk farms. The dominance of one farm type over the other varies across the regression coefficients. Therefore, the contributions of different factor inputs to generations of the outputs vary across the farming types.

The dynamic setting used for estimation of TFP allows us to calculate the derivatives of the regression coefficients with respect to time (Table 5). As Kumbhakar and Sun (2012) put it, these derivatives allow us to find out whether TC is biased towards certain inputs or outputs. The smooth coefficient model, in this case, produces the observation-specific measures of the bias in TC. It can be seen that TC is likely to slow down over time as the derivative associated with β_0 is slightly negative at the means and medians (Q2) for both farming types. As regards the input bias in TC, the divergent directions are observed across the five inputs. More specifically, TC for most of the farms are found to be labour-saving, given the significantly negative derivative at the median associated with β_1; and intermediate-consumption- and asset-using, given the significantly positive derivatives at the medians associated with β_3 and β_4, respectively. These suggest that traditional inputs are becoming relatively less important than those related to input-intensive technologies. The intermediate consumption and assets might become the limiting factors if farming revenue does not allow expanding their inflow. These effects are positive and significant at the 10% level at the medians for both types of farming.

These results can be indirectly compared with similar studies in other European countries. For instance, Mennig and Sauer (2019) reported land-saving, intermediate-consumption-saving, capital-using and livestock-using TC for German dairy farms during 2006–2011. Similar results were reported by Skevas, Emvalomatis, and Brümmer (2018b) for German dairy farms during 2001–2009. Compared with our result, the capital-using TC was obtained in both cases, yet intermediate-consumption-using TC is only observed for Lithuania. This suggests that modernisation of dairy farming involving more intensive farming practices are still important in Lithuania.

Finally, the signs of the derivatives associated with the output coefficients indicate scale bias in TC. As regards the milk output, derivative of the associated coefficient γ_1 is not significant at the median for milk farms, yet it is significantly negative at the median for milk-cattle farms. This indicates that most milk-cattle farms are operating above their efficient scale for producing milk. As regards derivative of γ_2 at the median, no significant conclusions can be identified for the milk farms, whereas most milk-cattle farms are operating below their efficient scale for producing other outputs.

5.2.2. Decomposition of TFP growth

To examine the temporal behaviour of TFP growth and its components, Fig. 4 plots the weighted average of TFP growth (Baltagi & Griffin, 1988) and its components over time where the weights are taken from the FADN.[25] These growth rates are used to define their respective indices from $I_t = I_{t-1}(1 + \dot{I}_t)$, where I is either TFP or one of its components, and $I_{2004} = 100$. A positive (nega-

[25] The FADN involves a representative survey based on the random sampling. Thus, each observation represents a certain number of farms in Lithuania. We use these numbers (normalised by their sum) as weights for constructing the weighted averages. Weights are used in this section instead of other sections for ease of plotting TFP growth and its components over time. Note that the plots are similar if we use simple averages instead.

Table 6
Summary statistics of TFP growth and its components.

	TC	Scale	Allocative	External	EC	Residual	TFP Growth
Full sample							
Mean	0.0189	0.0181	−0.0158	−0.0002	−0.0011	0.0005	0.0203
	(0.0009)	(0.0023)	(0.0023)	(0.0006)	(0.0000)	(0.0029)	(0.0046)
Q1	−0.0116	−0.0299	−0.0613	−0.0135	−0.0014	−0.0760	−0.1011
	(0.0004)	(0.0014)	(0.0015)	(0.0004)	(0.0001)	(0.0023)	(0.0033)
Q2	0.0306	0.0050	−0.0072	−0.0004	−0.0009	0.0004	0.0199
	(0.0019)	(0.0009)	(0.0015)	(0.0004)	(0.0000)	(0.0017)	(0.0040)
Q3	0.0460	0.0562	0.0433	0.0130	−0.0006	0.0776	0.1430
	(0.0007)	(0.0016)	(0.0017)	(0.0004)	(0.0000)	(0.0022)	(0.0041)
Type = Milk							
Mean	0.0182	0.0140	−0.0173	0.0007	−0.0011	0.0031	0.0176
	(0.0008)	(0.0021)	(0.0025)	(0.0006)	(0.0000)	(0.0029)	(0.0045)
Q1	−0.0114	−0.0277	−0.0616	−0.0120	−0.0014	−0.0720	−0.1008
	(0.0005)	(0.0016)	(0.0018)	(0.0004)	(0.0001)	(0.0024)	(0.0033)
Q2	0.0306	0.0039	−0.0087	0.0006	−0.0009	0.0028	0.0185
	(0.0021)	(0.0008)	(0.0015)	(0.0004)	(0.0000)	(0.0019)	(0.0040)
Q3	0.0452	0.0506	0.0403	0.0133	−0.0006	0.0776	0.1383
	(0.0004)	(0.0015)	(0.0016)	(0.0004)	(0.0000)	(0.0023)	(0.0039)
Type = Milk Cattle							
Mean	0.0251	0.0582	−0.0011	−0.0097	−0.0010	−0.0243	0.0471
	(0.0040)	(0.0150)	(0.0090)	(0.0028)	(0.0001)	(0.0126)	(0.0220)
Q1	−0.0153	−0.0605	−0.0442	−0.0247	−0.0012	−0.1279	−0.1135
	(0.0011)	(0.0062)	(0.0061)	(0.0023)	(0.0004)	(0.0137)	(0.0197)
Q2	0.0350	0.0329	0.0104	−0.0102	−0.0008	−0.0355	0.0403
	(0.0039)	(0.0077)	(0.0053)	(0.0012)	(0.0000)	(0.0085)	(0.0161)
Q3	0.0547	0.1416	0.0725	0.0065	−0.0006	0.0758	0.2135
	(0.0054)	(0.0106)	(0.0103)	(0.0019)	(0.0000)	(0.0076)	(0.0220)

Bootstrapped standard errors are in the parentheses.

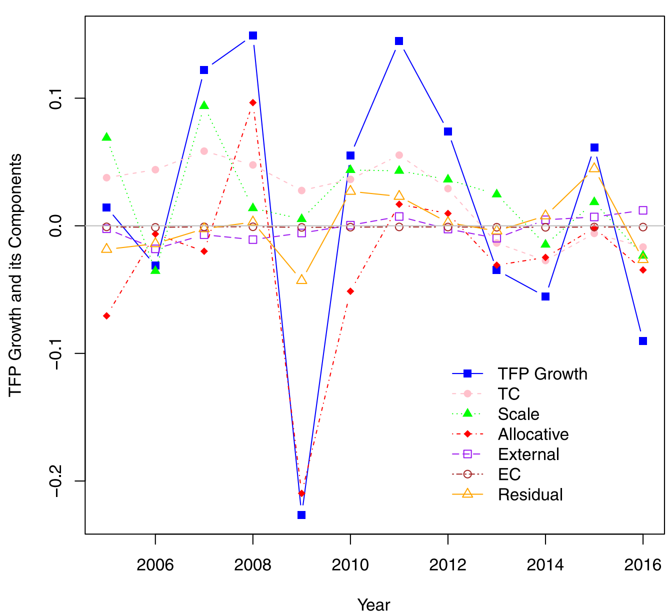

Fig. 4. TFP growth and its components.

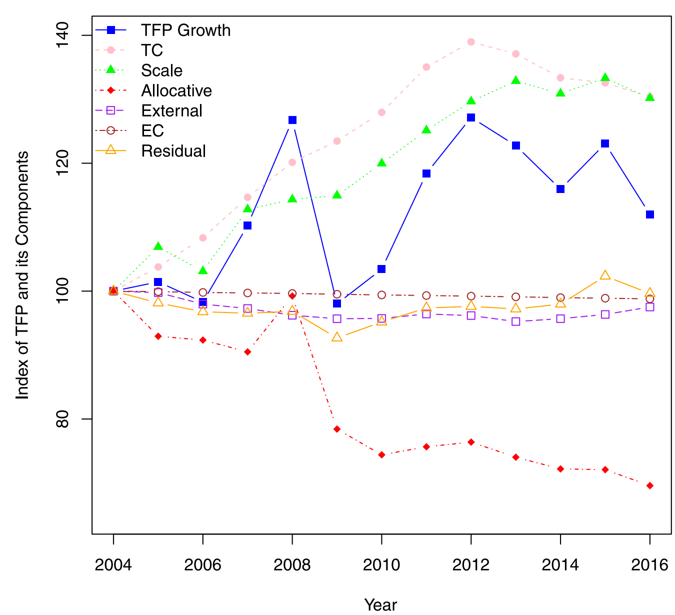

Fig. 5. TFP index and its components.

tive) growth rate in year t indicates that the corresponding index would rise (fall) from year $t-1$ to year t. For example, TFP growth is rather negative in year 2009 due to the EU dairy market crisis, and therefore the TFP index falls in year 2009 based on the previous year. Thus, these indices reveal the temporal behavior of TFP growth and its components.

In Fig. 5, TFP index suggests an appreciable productivity growth from the years 2006 to 2008. Then there is a sharp decline of productivity growth in 2009 before gradual recovery. The TFP index shows a more uneven trend than the other indices. It is clear from this figure that (1) the TC and scale components make the largest positive contribution to TFP growth, (2) the effects of EC, and external and residual components are relatively negligible, and (3) the allocative component contributes negatively to TFP growth. These results confirm the summary statistics reported in Table 6. Note that the values of TFP growth and its components vary across the quartiles so that the first and third ones often show different signs. Therefore, Lithuanian dairy farms are diverse in terms of the underlying TFP growth paths.

The slight decline in TFP index in 2006 can be related to the drought and the subsequent decrease in the feed supply. Year 2009 marked the turmoil in the EU dairy market amidst the global economic crisis. Similar effects were also reported by Mennig and Sauer (2019). This causes decrease in milk prices with subsequent reduction of milk supply and herd size, and leads to allocative inefficiency as evidenced by large negative shift in the allocative component of TFP growth in Fig. 4 and TFP index in Fig. 5. The decline in TFP post-2015 can be related to the Russian embargo which induces similar processes as in 2009. The decrease in the herd size affects the input structure and may have negative effect in the subsequent time periods due to inefficient use of the remaining resources and adjustment costs. The scale component remains the driving force, whereas the TC shows slight decline post-2012.

6. Conclusions

In this paper, we consider the semi/nonparametric analysis of technical inefficiency and TFP growth. To facilitate the analysis, we apply the general additive and SF models to extract the persistent and transient technical inefficiencies under the unknown functional form of the IDF. We allow both the mean and standard deviation of the transient inefficiency distribution to be dependent on the contextual variables. The smooth coefficient model is then applied in a dynamic setting to estimate and decompose TFP growth. The proposed methodology renders a highly flexible representation of the production technology and meaningful regression coefficients that can be interpreted in line with economic theory.

The empirical application focuses on Lithuanian dairy farms. More specifically, specialised milk farms and mixed milk-cattle farms are considered in the analysis. The performance gap between the two types of farms persists during 2004–2016, yet its magnitude is rather limited as the mean TTE score is 0.95, whereas the mean PTE score is 0.97.

The estimation of the smooth coefficient model sheds some light on the underlying technology and TFP growth. First, that the TC biases towards intermediate consumption and assets indicates the increasing spread of more intensive farming requiring higher level of modernisation. Lithuanian dairy farms maintained TFP growth of 2% per annum on average during 2004–2016, and much of it is attributed to the TC and scale components. However, TC shows a downward trend post-2012 which calls for technical innovations in the sector. Among other reasons, the input price growth rates exceeding those of the output prices may explain slowdown of the TC when monetary variables are involved in the analysis (Kumbhakar, Lien, Flaten, & Tveterås, 2008). The scale component indicates that higher market integration is likely to help Lithuanian farms to further improve their TFP by adjusting their production structure. Therefore, both technological and managerial innovations are needed to maintain the growth of TFP amidst the economic fluctuations rendered by the geopolitical conditions (e.g., Russian embargo; see Kutlina-Dimitrova, 2017) in Central and Eastern Europe.

The average annual growth in TFP for Lithuanian dairy farms obtained in this study is comparable to similar research in the other European countries. For instance, Mennig and Sauer (2019) applied SF analysis and obtained a TFP growth rate of 2.2% for German dairy farms. Skevas et al. (2018b) applied several SF models and obtained annual growth rates of around 1.7% (with additional case of -1% when trend is assumed to impact the inefficiency). In these two studies, TC appeared as the prevailing component behind TFP growth. Sipiläinen et al. (2013) reported 2.8% for Finnish dairy farms. Turning to nonparametric analysis, Madau, Furesi, and Pulina (2017) reported an average annual TFP growth rate of about 1% for the EU dairy farms. Depending on the farm size, Keizer and Emvalomatis (2014) obtained the average annual rates of TFP growth bounded between 0.4% and 2% for Dutch dairy farms.

We observe positive impacts of diversification on TC, one of the most important sources of TFP growth. This implies that public support (e.g., the CAP measures) should seek to promote diversification of production. To this end, measures for modernisation (i.e., investment support) and supply chain improvement (e.g., Rural Development Programme measures) can be designed and implemented. Indeed, these measures allow farmers to expand the scope of production by acquiring new equipment. Furthermore, they can embark on direct sales of processed products produced on farm, and thus generate higher profit margins (Ventura & Milone, 2000).

Future studies could include estimation of alternative representations of production technology within the current econometric modelling framework. Alternatively, different farming types may be covered to gain more insights into the possible development paths for the agricultural sector.

Acknowledgments

The authors would like to thank three anonymous referees for insightful comments, and are responsible for all remaining errors. This research is funded by (1) the European Social Fund (project No 09.3.3-LMT-K-712-01-0007) under grant agreement with the Research Council of Lithuania (LMTLT); and (2) the National Natural Science Foundation of China (Grant ID number: 71801146).

Appendix A. Alternative specifications of the determinants of transient inefficiency

In addition to the baseline results of determinants of the transient inefficiency reported in Table 3, we also provide in Table A.1 and Fig. A.1 the estimation results of two alternative specifications of the determinants of transient inefficiency. The first alternative specification controls for farm type, and the second alternative controls for farm type as well as regional dummies representing the ten counties in Lithuania. It can be seen that adding these additional dummies into $\mu(Z, t)$ and $\sigma_u(Z, t)$ does not change the signs of the variables used in the baseline specification, i.e., HHI, support payments, farmer's age, and time trend. Furthermore, the associated transient TE scores from the baseline specification in Fig. 1 are quite close to those from the alternative specifications

Table A.1
Determinants of the transient inefficiency.

	Farm type effects		Region effects	
	Coef.	Std. Err.	Coef.	Std. Err.
$\mu(Z, t)$				
Intercept	−0.0570	0.1016	−0.0588*	0.0357
HHI	−0.0196	0.0406	−0.0135	0.0125
Support payments	0.0046	0.0100	0.0065*	0.0034
Farmer's age	0.0005*	0.0003	0.0001	0.0002
Time trend	0.0005	0.0021	0.0009	0.0009
$\sigma_u(Z, t)$				
Intercept	−3.6559***	0.5659	−3.4335***	0.5138
HHI	−0.1549	0.2788	−0.0865	0.2521
Support payments	0.1290**	0.0559	0.1329***	0.0493
Farmer's age	−0.0075**	0.0038	−0.0080**	0.0039
Time trend	0.0255*	0.0146	0.0212	0.0142

1. Z variables are described in Section 4. 2. *, ** and *** indicate significance at the level of 10%, 5% and 1%, respectively.

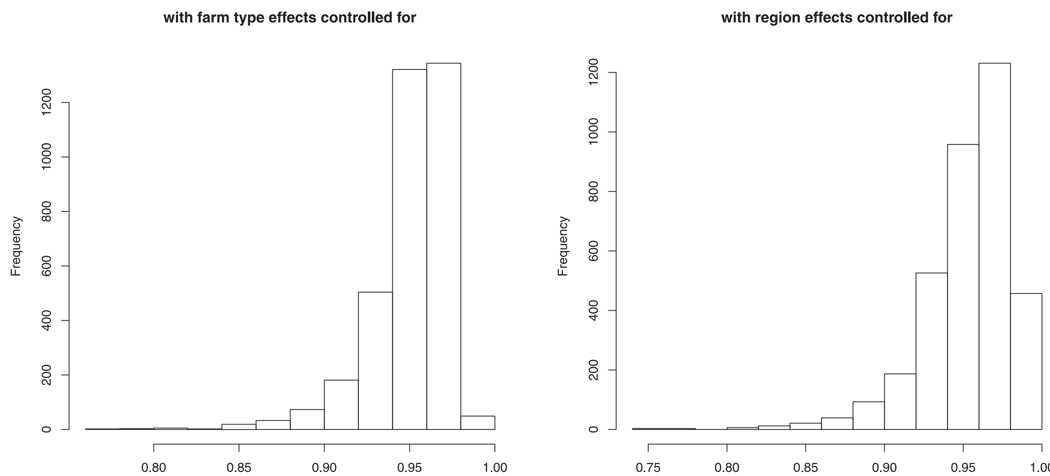

Fig. A.1. Transient TE scores under alternative specifications.

Table B.1
Determinants of the transient inefficiency.

	Level		Per hectare	
	Coef.	Std. Err.	Coef.	Std. Err.
$\mu(Z, t)$				
Intercept	−0.0115	0.0121	−0.0031	0.0168
HHI	−0.0032	0.0189	0.0149	0.0150
Support payments	0.0001	0.0004	−0.0001***	0.00004
Farmer's age	0.0002	0.0002	0.0005***	0.0001
Time trend	0.00004	0.0002	0.0048***	0.0008
$\sigma_u(Z, t)$				
Intercept	0.0026	7.5627	−0.3308	0.6948
HHI	0.0050	4.0467	−0.6054	0.4815
Support payments	−0.0494	0.4183	0.0202***	0.0012
Farmer's age	−0.1195**	0.0587	−0.0738***	0.0073
Time trend	−0.0276	0.0967	−0.6594***	0.0499

1. Farm type and regional dummies are included. 2. In the level model, support payments are measured in thousand EUR. 3. *, ** and *** indicate significance at the level of 10%, 5% and 1%, respectively.

in Fig. A.1; that is, they are all slightly skewed to the left, with a mean of 0.95.

Appendix B. Estimation of (6) with alternative measures of support payments

In addition to the baseline results of determinants of the transient inefficiency reported in Table 3, we also provide in Table B.1 the estimation results using two alternative measures of subsidies, i.e., support payments in level, and support payments per hectare. In the level model, the coefficient of support payments in $\mu(Z, t)$ is positive, but both coefficients of support payments in $\mu(Z, t)$ and $\sigma_u(Z, t)$ are insignificant at the 10% level. In the per hectare model, the coefficient of support payments in $\sigma_u(Z, t)$ is positive, and both coefficients of support payments in $\mu(Z, t)$ and $\sigma_u(Z, t)$ are significant at the 1% level.

Appendix C. Descriptive statistics of the variables for estimating (19)

In (19), growth rates are computed via first-order log differences. Those farms observed for only one year, therefore, are removed. Table C.1 presents the descriptive statistics of the variables used to estimate (19).

Table C.1
Descriptive statistics of the variables for estimating (19).

	Full sample	Type = Milk	Type = Milk Cattle
	Labour (X_1), hours		
Mean	6365.01	6485.64	5195.88
SD	5844.36	6070.35	2569.47
Min	1350.00	1350.00	2030.00
Max	93544.00	93544.00	14704.00
	Herd size (X_2), LSU		
Mean	67.07	67.71	60.88
SD	79.34	81.19	58.30
Min	2.75	2.75	4.62
Max	800.77	800.77	269.46
	Intermediate consumption (X_3), index		
Mean	59199.71	61110.72	40678.46
SD	95974.48	99543.84	45650.56
Min	2808.51	2808.51	4742.02
Max	1684415.29	1684415.29	242979.70
	Assets (X_4), index		
Mean	92141.13	95264.28	61871.94
SD	136481.90	140819.72	77441.88
Min	1333.69	1333.69	2314.96
Max	1150755.95	1150755.95	510487.32
	Utilised agricultural area (X_5), hectare		
Mean	100.20	100.72	95.15
SD	104.25	106.51	79.07
Min	2.17	2.17	4.40
Max	956.77	956.77	386.95
	Milk output (Y_1), ton		
Mean	236.87	251.83	91.88
SD	345.82	357.35	133.94
Min	2.16	7.53	2.16
Max	4867.88	4867.88	704.50
	Other outputs (Y_2), index		
Mean	44901.00	45020.00	43747.68
SD	71644.37	73654.61	48111.12
Min	98.19	98.19	4220.02
Max	1004395.41	1004395.41	262848.28
	Labour price (W_1), EUR/hour		
Mean	2.16	2.17	2.09
SD	0.61	0.62	0.52
Min	0.89	0.89	0.95
Max	6.76	6.76	3.64
	Herd price (W_2), EUR/LSU		
Mean	514.69	526.11	403.97
SD	188.64	190.93	117.06
Min	129.96	146.61	129.96
Max	2148.36	2148.36	811.24
	Intermediate consumption price (W_3), index		
Mean	1.09	1.09	1.08

(*continued on next page*)

Table C.1 (*continued*)

	Full sample	Type = Milk	Type = Milk Cattle
SD	0.14	0.14	0.15
Min	0.77	0.77	0.77
Max	1.26	1.26	1.26
	Assets price (W_4), EUR		
Mean	0.26	0.26	0.28
SD	0.12	0.12	0.12
Min	0.03	0.03	0.06
Max	0.63	0.62	0.63
	Land price (W_5), EUR/hectare		
Mean	15.10	15.28	13.38
SD	13.61	13.72	12.39
Min	0.05	0.05	0.15
Max	126.41	126.41	56.96
	Milk price (P_1), EUR/ton		
Mean	240.14	242.73	215.04
SD	56.28	55.77	55.20
Min	86.89	86.89	110.05
Max	434.43	434.43	367.82
	Other outputs' price (P_2), index		
Mean	1.00	1.00	0.99
SD	0.03	0.04	0.02
Min	0.93	0.93	0.95
Max	2.41	2.41	1.06
	HHI (Z_1)		
Mean	0.68	0.70	0.52
SD	0.20	0.19	0.19
Min	0.17	0.21	0.17
Max	1.00	1.00	1.00
	Support payments in log (Z_2)		
Mean	9.53	9.51	9.69
SD	1.03	1.04	0.94
Min	6.34	6.34	7.36
Max	12.59	12.59	11.60
	Farmer's age (Z_3), years		
Mean	47.88	48.05	46.26
SD	9.95	9.92	10.19
Min	21.00	22.00	21.00
Max	75.00	75.00	74.00
	Time trend (t)		
Mean	8.43	8.47	8.07
SD	3.14	3.13	3.19
Min	2.00	2.00	2.00
Max	13.00	13.00	13.00

1. Total number of observations = 1978. 2. t is defined as year-2003, where year varies from 2005 to 2016.

References

Aigner, D., Lovell, C. K., & Schmidt, P. (1977). Formulation and estimation of stochastic frontier production function models. *Journal of Econometrics, 6*(1), 21–37.

Ang, F., & Oude Lansink, A. (2017). Decomposing dynamic profit inefficiency of Belgian dairy farms. *European Review of Agricultural Economics, 45*(1), 81–99.

Badunenko, O., & Kumbhakar, S. (2017). Economies of scale, technical change and persistent and time-varying cost efficiency in Indian banking: Do ownership, regulation and heterogeneity matter? *European Journal of Operational Research, 260*, 789–803.

Baležentis, T., Li, T., & Baležentis, A. (2015). The trends in efficiency of Lithuanian dairy farms: A semiparametric approach. *Management Theory and Studies for Rural Business and Infrastructure Development, 37*(2), 167–178.

Baltagi, B., & Griffin, J. (1988). A general index of technical change. *Journal of Political Economy, 96*(1), 20–41.

Battese, G. E., & Coelli, T. J. (1988). Prediction of firm-level technical efficiencies with a generalized frontier production function and panel data. *Journal of Econometrics, 38*(3), 387–399.

Cabrera, V. E., Solis, D., & Del Corral, J. (2010). Determinants of technical efficiency among dairy farms in Wisconsin. *Journal of Dairy Science, 93*(1), 387–393.

Cameron, A., & Trivedi, K. (2005). *Microeconometrics: methods and application* (pp. 376–377). Cambridge University Press, Chapter 11.

Cechura, L., Grau, A., Hockmann, H., Levkovych, I., & Kroupova, Z. (2017). Catching up or falling behind in European agriculture: The case of milk production. *Journal of Agricultural Economics, 68*(1), 206–227.

Chambers, R. (1988). *Production economics*. Cambridge University Press.

Coelli, T., & Perelman, S. (1999). A comparison of parametric and non-parametric distance functions: With application to European railways. *European Journal of Operations Research, 117*, 326–339.

Coelli, T., Rao, P., & Battese, G. (1998). *An introduction to efficiency and productivity analysis*. Kluwer Academic Publishers, Boston.

Colombi, R., Kumbhakar, S., Martini, G., & Vittadini, G. (2014). Closed-skew normality in stochastic frontiers with individual effects and long/short-run efficiency. *Journal of Productivity Analysis, 42*, 123–136.

Commission, E. (2019). Farm accountancy data network. https://ec.europa.eu/agriculture/rica/.

Diewert, W. (1981). *The economic theory of index numbers: A survey*. Cambridge University Press, London.

Economics, L. I. o. A. (2017). *FADN Survey Results 2016*. Lietuvos Agrarines Ekonomikos Institutas, Vilnius.

Fan, Y., Li, Q., & Weersink, A. (1996). Semiparametric estimation of stochastic production frontier models. *Journal of Business and Economic Statistics, 14*, 460–468.

Feng, G., & Serletis, A. (2010). Efficiency, technical change, and returns to scale in large US banks: Panel data evidence from an output distance function satisfying theoretical regularity. *Journal of Banking & Finance, 34*, 127–138.

Jondrow, J., Lovell, C. A. K., Materov, I. S., & Schmidt, P. (1982). On the estimation of technical inefficiency in the stochastic frontier production function model. *Journal of Econometrics, 19*, 233–238.

Keizer, T., & Emvalomatis, G. (2014). Differences in TFP growth among groups of dairy farms in the Netherlands. *NJAS-Wageningen Journal of Life Sciences, 70*, 33–38.

Kuipers, A. y., Malak-Rawlikowska, A., Stalgiene, A., & Klopcic, M. (2017). Analysis of stakeholders' expectations for dairy sector development strategies from a central-eastern and western European perspective. *German Journal of Agricultural Economics, 66*(4), 265–280.

Kumbhakar, S. (2013). Specification and estimation of multiple output technologies: A primal approach. *European Journal of Operational Research, 231*, 465–473.

Kumbhakar, S., & Heshmati, A. (1995). Efficiency measurement in Swedish dairy farms: An application of rotating panel data, 1976-88. *American Journal of Agricultural Economics, 77*, 660–674.

Kumbhakar, S., Lien, G., Flaten, O., & Tveterås, R. (2008). Impacts of Norwegian milk quotas on output growth: A modified distance function approach. *Journal of Agricultural Economics, 59*(2), 350–369.

Kumbhakar, S., Lien, G., & Hardaker, J. (2014). Technical efficiency in competing panel data models: A study of Norwegian grain farming. *Journal of Productivity Analysis, 41*, 321–337.

Kumbhakar, S., & Lovell, C. (2000). *Stochastic frontier analysis*. Cambridge University Press, Cambridge.

Kumbhakar, S., Park, B. U., Simar, L., & Tsionas, E. G. (2007). Nonparametric stochastic frontiers: A local maximum likelihood approach. *Journal of Econometrics, 137*(1), 1–27.

Kumbhakar, S., & Sun, K. (2012). Estimation of TFP growth: A semiparametric smooth coefficient approach. *Empirical Economics, 43*, 1–24.

Kumbhakar, S., & Sun, K. (2013). Derivation of marginal effects of determinants of technical inefficiency. *Economics Letters, 120*(2), 249–253.

Kutlina-Dimitrova, Z. (2017). The economic impact of the russian import ban: a CGE analysis. *International Economics and Economic Policy, 14*(4), 537–552.

Lai, H., & Kumbhakar, S. (2018). Panel data stochastic frontier model with determinants of persistent and transient inefficiency. *European Journal of Operational Research, 271*, 746–755.

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

Latruffe, L., Bravo-Ureta, B. E., Carpentier, A., Desjeux, Y., & Moreira, V. H. (2017). Subsidies and technical efficiency in agriculture: Evidence from European dairy farms. *American Journal of Agricultural Economics, 99*(3), 783–799.

de Lauwere, C., Malak-Rawlikowska, A., Stalgiene, A., Klopcic, M., & Kuipers, A. (2018). Entrepreneurship and competencies of dairy farmers in Lithuania, Poland and Slovenia. *Transformations in Business & Economics, 17*(3), 237–257.

Madau, F. A., Furesi, R., & Pulina, P. (2017). Technical efficiency and total factor productivity changes in European dairy farm sectors. *Agricultural and Food Economics, 5*(1), 17.

Mary, S. (2013). Assessing the impacts of pillar 1 and 2 subsidies on TFP in French crop farms. *Journal of Agricultural Economics, 64*(1), 133–144.

Meesters, A. (2014). A note on the assumed distributions in stochastic frontier models. *Journal of Productivity Analysis, 42*(2), 171–173.

Meeusen, W., & van Den Broeck, J. (1977). Efficiency estimation from Cobb-Douglas production functions with composed error. *International Economic Review*, 435–444.

Mennig, P., & Sauer, J. (2019). The impact of agri-environment schemes on farm productivity: A DID-matching approach. *European Review of Agricultural Economics*. doi:10.1093/erae/jbz006.

Minviel, J. J., & De Witte, K. (2017). The influence of public subsidies on farm technical efficiency: A robust conditional nonparametric approach. *European Journal of Operational Research, 259*(3), 1112–1120.

Minviel, J. J., & Latruffe, L. (2017). Effect of public subsidies on farm technical efficiency: A meta-analysis of empirical results. *Applied Economics, 49*(2), 213–226.

Restrepo-Tobon, D., Kumbhakar, S., & Sun, K. (2015). Obelix vs. asterix: Size of US commercial banks and its regulatory challenge. *Journal of Regulatory Economics, 48*(2), 125–168.

de Roest, K., Ferrari, P., & Knickel, K. (2018). Specialisation and economies of scale or diversification and economies of scope? assessing different agricultural development pathways. *Journal of Rural Studies, 59*, 222–231.

Shephard, R. (1953). *Cost and production functions*. Princeton University Press, Princeton.

Shephard, R. (1970). *Theory of cost and production functions*. Princeton University Press, Princeton.

Sipiläinen, T., Kumbhakar, S. C., & Lien, G. (2013). Performance of dairy farms in Finland and Norway from 1991 to 2008. *European Review of Agricultural Economics, 41*(1), 63–86.

Skevas, I., Emvalomatis, G., & Brümmer, B. (2017). The effect of farm characteristics on the persistence of technical inefficiency: A case study in German dairy farming. *European Review of Agricultural Economics, 45*(1), 3–25.

Skevas, I., Emvalomatis, G., & Brümmer, B. (2018a). Heterogeneity of long-run technical efficiency of German dairy farms: A Bayesian approach. *Journal of Agricultural Economics, 69*(1), 58–75.

Skevas, I., Emvalomatis, G., & Brümmer, B. (2018b). Productivity growth measurement and decomposition under a dynamic inefficiency specification: The case of German dairy farms. *European Journal of Operational Research, 271*(1), 250–261.

Sun, K., & Kumbhakar, S. C. (2013). Semiparametric smooth-coefficient stochastic frontier model. *Economic Letters, 120*(2), 305–309.

Sun, K., Kumbhakar, S. C., & Tveterås, R. (2015). Productivity and efficiency estimation: A semiparametric stochastic cost frontier approach. *European Journal of Operational Research, 245*(1), 194–202.

Tsionas, E., & Kumbhakar, S. (2014). Firm heterogeneity, persistent and transient technical inefficiency: A generalized true random-effects model. *Journal of Applied Econometrics, 29*, 110–132.

Ventura, F., & Milone, P. (2000). Theory and practice of multi-product farms: Farm butcheries in umbria. *Sociologia Ruralis, 40*(4), 452–465.

Verhees, F., Malak-Rawlikowska, A., Stalgiene, A., Kuipers, A., & Klopčič, M. (2018). Dairy farmers business strategies in Central and Eastern Europe based on evidence from Lithuania, Poland and Slovenia. *Italian Journal of Animal Science, 17*(3), 755–766.

Wang, H.-J., & Schmidt, P. (2002). One-step and two-step estimation of the effects of exogenous variables on technical efficiency levels. *Journal of Productivity Analysis, 18*, 129–144.

Yao, F., Zhang, F., & Kumbhakar, S. (2019). Semiparametric smooth coefficient stochastic frontier model with panel data. *Journal of Business & Economic Statistics, 37*(3), 556–572.