Computational Economics (2025) 66:4381–4416
https://doi.org/10.1007/s10614-025-10857-9

Skewness Issues in Quantifying Efficiency: Insights from Stochastic Frontier Panel Models Based on Closed Skew Normal Approximations

Rouven E. Haschka[1,2] · Dominik Wied[3]

Accepted: 7 January 2025 / Published online: 22 January 2025

Abstract
Typically, the inefficiency term in stochastic frontier models is assumed to be positively skewed; however, efficiency scores are biased if this assumption is violated. This paper considers the case in which negative skewness is also allowed in the model. The paper discusses estimation of a stochastic frontier panel model with unobserved fixed effects without having to identify additional parameters that determine skewness of inefficiency. On the one hand, the parameters can be estimated via integrating out nuisance parameters by means of marginal maximum likelihood. On the other hand, we propose an approximation based on closed skew normal distributions, which turns out to be sufficiently accurate for maximum likelihood estimation. Simulations assess the finite sample performance of estimators and show that all model parameters and efficiency scores can be estimated consistently regardless of positive or negative inefficiency skewness. An empirical analysis to unravel inefficiencies in the German healthcare system demonstrates the practical relevance of the model.

Keywords Fixed effects · Panel data · Skewness · Stochastic frontier analysis · Closed skew normal distribution

Mathematics Subject Classification C23 · D24 · I11 · I18

R. E. Haschka is affiliated with the Chair of Business Analytics & Data Science, Zeppelin University, and the Institute of Strategy and Management, Corvinus University. Dominik Wied is affiliated with the University of Cologne, Germany.

Extended author information available on the last page of the article

1 Introduction

The classical assumption in stochastic frontier (SF) models is that errors exhibit a negative skewness. However, in empirical applications, the residuals may have positive skewness. Waldman (1982) first demonstrated that if the residuals from the SF model have ‘wrong’ skewness, inefficiency variance is biased toward zero. In consequence, efficiency scores tend to be one, which leads to false conclusions of high efficiency (Hafner et al., 2018). Green and Mayes (1991) argue that it either indicates ‘super efficiency’ (all producers in the industry operate close to the frontier), or the inappropriateness of the technique of SF analysis to measure inefficiencies. Thus, implausibly high efficiency scores obtained under the classical SF specification can be an indicator of misspecifying inefficiency skewness (Hafner et al., 2018). The problem of wrongly skewed residuals in SF models has been recognized in various fields of application such as agricultural economics (Curtiss et al., 2021; Vargas-Leitón et al., 2015), tourism management (Choi et al., 2021), banking (Wheelock & Wilson, 2020), and macroeconomics (Daniel et al., 2019). Green and Mayes (1991) report that for a sample of 151 UK industries, 32% showed wrong residual skewness, and for a sample of 1401 Australian industries, a similar problem was encountered in 35% of the cases. Mester (1997) estimates a cost frontier and finds that out of 12 US bank districts, three have wrongly skewed residuals.

According to Almanidis and Sickles (2011) and Hafner et al. (2018), the most likely reason for detecting wrong skewness is small samples. Simar and Wilson (2010) confirms this, highlighting that small samples can lead to wrong skewness despite correct skewness in the underlying population. Apart from sample issues, the reasons for the occurrence of wrong skewness are questioned remarkably little (for an excellent recent review, see Papadopoulos and Parmeter, 2023). On the one hand, the methodological literature dealing with the consequences of wrong skewness figured out (i) asymmetry of idiosyncratic noise, (ii) dependence between inefficiency and idiosyncratic noise, and (iii) wrong skewness of inefficiency as potential reasons (Haschka, 2024b). On the other hand, economic explanatory approaches for wrong skewness deal with the issue in general (Papadopoulos & Parmeter, 2023), but do not delve into investigating where (i.e., from which component) it might originate.

The access to panel data enables estimating inefficiency with greater consistency in comparison with cross-sectional data, because panel estimators strongly benefit from increases in the time dimension (Jondrow et al., 1982; Wang & Ho, 2010). Extant studies considering wrong skewness have only dealt with cross-sectional models, but not yet with panel SF models. Given this background, this study proposes a panel data generalization with fixed effects of the model by Hafner et al. (2018) that is able to distinguish correct and wrong skewness in one model without imposing a priori sign restrictions on inefficiency skewness. In this respect, we follow the literature about point (iii) by assuming that wrong skewness originates from inefficiency. We propose a precise approximation of the density of composed errors by means of the closed skew normal (CSN) distribution.

As properties of the CSN distribution can then be exploited, the generalization to fixed-effects transformed panel models is uncomplicated. Our model nests that of Chen et al. (2014), but it should be mentioned that it is difficult to consider external determinants of inefficiency in the proposed approach because numerical optimization would be cumbersome in such a case. We conduct Monte Carlo experiments and assess the performance of the proposed generalization. We further consider marginal maximum likelihood estimation with integrating out the individual-specific effects as a benchmark to assess the validity of the proposed CSN approximation. The simulations indicate that the estimator (i) remains consistent regardless of whether the distribution of inefficiency is positively or negatively skewed, (ii) can be applied to small sample sizes, and (iii) experiences minimal performance losses due to approximation.

The empirical analysis aims for an unbiased assessment of inefficiencies in the German healthcare system using a panel dataset on district-specific health profiles (data was taken from Herwartz and Schley, 2018). We detect wrong skewness in ten out of sixteen federal states in Germany. Accordingly, it exposes extant studies using traditional SF approaches to criticism for biased efficiency estimation. For instance, Herwartz and Schley (2018) assume correct skewness and report rather high efficiency scores for Eastern Germany (above.99 on average). However, our results show that wrong skewness is actually present for this region, which could be responsible for such high efficiencies. By contrast, the proposed estimator yields efficiency scores around.71 on average for Eastern Germany. As a policy recommendation, wrong skewness in health care markets (which are not subject to competitive conditions) means that additional interventions for health policy are needed to create incentives for efficiency improvements.

The paper begins by reviewing the current literature discussion about the sources of wrong skewness, followed by a brief review of the panel SF literature. Subsequently, we discuss how to consider both correct and wrong skewness in panel SF modelling with fixed effects. We then characterise the finite sample performance of the estimators using several Monte Carlo studies. Finally, we demonstrate the applicability of the proposed estimators to real data.

2 Literature Reviews

We first discuss the reasons for wrong skewness considered in the literature. Subsequently, we review methodological developments in panel stochastic frontier modeling.

2.1 Reasons for 'Wrong' Skewness

Stochastic frontier models often assume positive skewness in the inefficiency distribution to reflect competitive market dynamics, where producers operate (Aigner et al., 1977). This assumption rests on economic theory, particularly in microeconomic production models that posit producers aim to optimize output given their inputs. In competitive markets, producers minimize costs and maximize outputs, thus operating close to the efficiency frontier, under pressure to improve efficiency. Inefficiency appears as deviations from this frontier, with highly inefficient producers likely exiting the market (Haschka & Herwartz, 2020, 2022). Therefore, specifying stochastic frontier models with positively skewed inefficiency distributions, such as the common half-normal distribution (Kumbhakar et al., 2020), is appropriate for evaluating producer efficiency in competitive markets.

The prevalent view attributes the detection of wrong skewness, often seen as an empirical phenomenon (Almanidis & Sickles, 2011; Hafner et al., 2018; Waldman, 1982), to small sample sizes (Simar and Wilson, 2010, pp.8–9). While insufficient sample size is commonly blamed for cases where observed skewness deviates from true skewness, other explanations remain underexplored (Haschka, 2024b; Papadopoulos & Parmeter, 2023). Although some studies have identified cases of wrong skewness in empirical applications (e.g., Almanidis and Sickles,2011, Hafner et al., 2018, Haschka, 2024a, Parmeter and Racine, 2013), they often lack a thorough examination of potential population characteristics contributing to this phenomenon.

Certain data structure characteristics may contribute to wrong skewness. The asymmetry of the idiosyncratic error term can induce multimodality in efficiency score distributions, thus leading to wrong skewness (e.g., Badunenko and Henderson, 2023, Bonanno et al., 2017, Horrace et al., 2023, Son et al., 1993). Additionally, unmodelled dependence between idiosyncratic noise and inefficiency can also be a source (e.g., Smith, 2008, Bonanno et al., 2017, Bonanno and Domma, 2022). From an economic perspective, specific population characteristics, such as unique market features, can explain wrong skewness. Carree (2002) suggests industries characterized by cycles of innovation and imitation, where few firms innovate while many remain inefficient, resulting in wrong skewness. Akio (1992) highlights how technological progress and slow asset replacement among producers lead to misalignment, as another economic explanation of wrong skewness. These explanations imply that the inefficiency term is responsible for skewness issues, although not explicitly labeled as such. In such scenarios, producers might not face adequate pressure to maximize efficiency (Haschka, 2024b). Markets characterized by heavy regulation or entry barriers could reinforce this (Papadopoulos & Parmeter, 2023). Additionally, factors like limited market transparency (Møllgaard & Overgaard, 2001), technological constraints (Ortega, 2010), market imperfections (Cohen & Winn, 2007), seller's markets (Redmond, 2013), or structural factors (Haschka, 2024a) could further dampen competitive forces. In the absence of alternative explanations for wrong skewness, such as poor samples, the prevalence of producers operating at lower efficiency levels, with only a minority nearing the frontier, may stem from limited competition in the market.

Upon reviewing these contributions, two key observations emerge. Firstly, the literature addressing methodological explanations for wrong skewness overlooks the linking of these reasons to population characteristics. Understanding market mechanisms that introduce dependence between idiosyncratic noise and inefficiency or cause asymmetry in noise distribution necessitates sound economic explanations. For instance, explaining the simultaneous increase (or decrease) in efficiency due to unobserved production shocks and the prevalence of positive shocks over negative ones (or vice versa) requires deeper exploration. Secondly, when addressing skewness issues from an economic perspective, determining which model components need adjustment to reflect this peculiarity becomes crucial. Beyond examining dependence within composite errors or asymmetry of idiosyncratic noise, attributing skewness issues to the inefficiency term offers valuable insights into general competitive market dynamics (Haschka, 2024b). In light of this, we follow the literature associating wrong skewness with the inefficiency term in our methodological contribution.

2.2 Recent Developments in Panel Data SF Modeling

The advantage of panel data is that more information on inefficiency and productivity can be parsed, and in particular, shed light on changes in efficiency or productivity, which differs from a cross-sectional setting that can only provide a static portrayal of inefficiency. Furthermore, panel data models allow separating the effects of differences across individuals from effects that vary over time within individuals. Earlier work on panel data SF modeling treats unobserved heterogeneity as part of inefficiency (Schmidt & Sickles, 1984; Battese & Coelli, 1992). Such time-invariant SF models allow inefficiency to differ across individuals but restrict any change over time. The implication of this is that an inefficient producer cannot improve productivity over time by lessening inefficiency, which might be unrealistic when the time dimension is large (Kumbhakar et al., 2014). Furthermore, if producer-specific heterogeneity is not adequately controlled for, estimated inefficiency might pick up producer-specific heterogeneity in addition to, or even instead of inefficiency.

Greene (2005) proposes a standard fixed-effects panel model that is augmented by a stochastic inefficiency term. In this so-called ‘true fixed-effects model’, heterogeneity is captured by dummy variables to account for unobserved heterogeneity. However, the difficulty of this model is the incidental parameters problem. In samples with small time dimensions, there is a noticeable bias in both parameter estimates and efficiency. To avoid the need for dummy variables, Wang and Ho (2010) use fixed-effects transformation to eliminate the individual fixed effects; but their model requires that stochastic inefficiency be time-invariant and depend on observed time-varying regressors. Belotti and Ilardi (2012) reconsider Greene’s model and also use fixed-effects transformation, but allow for time-varying stochastic inefficiency. Since the model transformation does not lead to a known distribution for one-sided inefficiency, the authors use a marginal maximum simulated likelihood approach and integrates the inefficiency term out of the model. However, the approach requires inefficiency to depend on time-invariant explanatory variables.

Domınguez-Molina et al. (2003) recognizes the relevance of the closed-skew normal distribution (CSN) in SF modeling. More precisely, when assuming normally distributed idiosyncratic noise coupled with half normally distributed inefficiency, the composed error follows a CSN distribution (González-Farıas et al., 2004). Chen et al. (2014) adopt Greene’s model, use fixed-effects transformation, and show that the transformed errors again have a CSN distribution that can be exploited for maximum likelihood estimation. Belotti and Ilardi (2018) generalize the model of Chen et al. (2014)

by allowing for heteroskedastic and autocorrelated inefficiency, and the modeling of inefficiency variance as a function of exogenous effects.

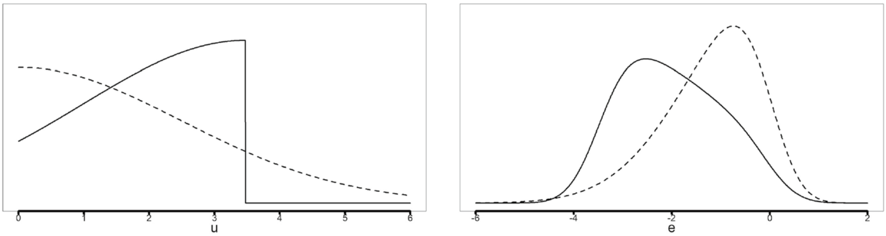
Although exploiting properties of the CSN distribution became a pertinent approach in panel SF modeling, all studies to date only consider positively skewed inefficiency by predominantly assuming a half-normal distribution. In such a specification, it is inherently assumed that the majority of producers operate close to full efficiency, which seems reasonable in competitive markets. As already mentioned, a lack of competition in the market likely supports the assumption of negative skewness if a major share of producers displays significant levels of inefficiency; however, this phenomenon does not necessarily require inefficiency to have negative skewness. The consideration of the inefficiency distribution with mode bounded away from zero could accommodate it, such as the truncated normal distribution when introducing a (strictly) positive location parameter. Accordingly, both distributions with positive and negative skewness could map inefficiencies in uncompetitive markets. By contrast, the advantage of using a negatively skewed distribution for inefficiency is that it can be specified relatively simply by using a truncated normal distribution where the upper bound corresponds to the truncation point (Almanidis et al., 2014). When the upper bound and truncation point further depend on the variance parameter, the expectation of the truncated normal distribution is equal to that of the standard half-normal distribution with mode at zero (Hafner et al., 2018). This yields negatively skewed inefficiency without requiring an additional location parameter for identification. Recently, El Mehdi and Hafner (2024) generalize the model by Hafner et al. (2018) to the panel case and allow for time-decaying inefficiency in the spirit of Battese and Coelli (1992). However, they consider a pooled model and thus do not account for producer-specific heterogeneity.

Taken together, on the one hand, while there are several contributions to panel SF models considering producer heterogeneity (e.g., Belotti and Ilardi, 2018, Chen et al., 2014, Wang and Ho, 2010), negatively skewed inefficiency has not yet been considered. On the other hand, it remains questionable if the composed errors are still CSN distributed if inefficiency exhibits a truncated normal distribution with negative skewness. Against this background, this study is the first to propose a panel model with individual-specific fixed effects that is able to distinguish correct and wrong skewness. For this purpose, the panel model is subjected to fixed-effects transformation while exploiting properties of the CSN distribution enables the derivation of the likelihood function in closed form. We also provide theoretical results for the CSN distribution.

3 Methodology

Consider a typical panel data SF model for $i = 1, \dots, N$ cross-sectional units over time instances $t = 1, \dots, T$:

$$y_{it} = \alpha_i + \boldsymbol{x}'_{it}\boldsymbol{\beta} + \underbrace{v_{it} - u_{it}}_{e_{it}}, \tag{1}$$



(a) Density of u. (b) Density of e.

Fig. 1 Densities of u and e for $\gamma = 2.5$, i.e., correct skewness (dotted lines) and $\gamma = -2.5$, i.e., wrong skewness (solid lines); with $\sigma_v = .5$

where y_{it} is the dependent variable, α_i is panel i's (unobserved) time-invariant effect, and $\boldsymbol{xit}$ is a $(1 \times K)$ vector of regressors. The parameter vector $\boldsymbol{\beta}$ describes the linear influence of $\boldsymbol{xit}$ on y_{it}. By means of a fixed-effects treatment, α_i is allowed to be correlated with $\boldsymbol{x}_{it}$ and e_{it}.[1]

According to the SF specification, $v_{it} \sim N(0, \sigma_v^2)$ captures two-sided idiosyncratic noise and u_{it} is a truncated normally distributed inefficiency component on the positive scale. Following Hafner et al. (2018), we distinguish two cases for u_{it} which characterize the shape of composed errors $e_{it} = v_{it} - u_{it}$. The assumption $u_{it} \sim N_{[0,\infty)}(0, \gamma^2)$ with $\gamma > 0$ is well-disseminated and describes correct skewness of both u_{it} and e_{it} because the density of u_{it} is strictly decreasing in $[0, \infty)$ (Kumbhakar & Lovell, 2003).[2] In contrast, wrong skewness is induced by assuming $u_{it} \sim N_{[0,a_0|\gamma|)}(a_0|\gamma|, \gamma^2)$ with $\gamma < 0$. Here, $a_0 \approx 1.389$ is the non-trivial solution of $\frac{\phi(0)}{\Phi(0)} = a_0 + \frac{\phi(a_0)-\phi(0)}{\Phi(a_0)-\Phi(0)}$ (for the cdf Φ and the pdf ϕ of the standard normal distribution) and the density of u_{it} is strictly increasing and bounded in $[0, a_0|\gamma|]$. This is illustrated in Panel (a) of Fig. 1. It is worth highlighting that expectations of both u_{it} and e_{it} remain unaffected by the choice of distribution for u_{it} (Hafner et al., 2018). Although the sign of γ determines either positive or negative skewness, it does not affect expectations of u_{it} and e_{it}. Thus, inefficiency variance and sign of skewness are directly related because $\gamma > 0$ ($\gamma < 0$) induces correct (wrong) skewness but $\mathbb{E}[u_{it}]$ and $\mathbb{E}[e_{it}]$ are only subject to $|\gamma|$. The consideration of both correct and wrong skewness allows parameters of the stochastic frontier model to be identified regardless of the sign of the residuals (Hafner et al., 2018).

[1] It is important to note that in our model, the parameter α_i solely represents unobserved heterogeneity, which is not associated with inefficiency. However, there is literature that decomposes α_i into two distinct components: heterogeneity and persistent inefficiency (Colombi et al., 2014; Kumbhakar et al., 2014) While this alternative provides greater flexibility, it relies on strong identifying assumptions, such as a random effects treatment of both unobserved heterogeneity and time-invariant (persistent) inefficiency (Filippini & Greene, 2016). Consequently, it is necessary for these components to be independent from all other elements of the model, including regressors. In contrast, our approach permits correlation between α_i and both $\boldsymbol{x}_{it}$ and e_{it}.

[2] In general, correct skewness means positive skewness of u_{it} and in consequence negative skewness of $e_{it} = v_{it} - u_{it}$ due to symmetry of v_{it}.

Assuming that v_{it} and u_{it} are distributed independently from each other, the density of e_{it} is given by:

$$'Correct'\ skewness\ :\ f_e^+(e) = \frac{2}{\sigma}\phi\left(\frac{e}{\sigma}\right)\Phi\left(-\frac{e\gamma}{\sigma\sigma_v}\right), \tag{2}$$

$$\begin{aligned} &'Wrong'\ skewness: \\ &f_e^-(e) = \frac{1}{\sigma\left(\Phi\left(a_0\right) - \Phi(0)\right)}\phi\left(\frac{e - a_0\gamma}{\sigma}\right)\left[\Phi\left(A_w + \frac{a_0\sigma}{\sigma_v}\right) - \Phi\left(A_w\right)\right], \\ &A_w = \frac{e - a_0\gamma}{\sigma}\frac{\gamma}{\sigma_v}, \end{aligned} \tag{3}$$

with $\sigma^2 = \gamma^2 + \sigma_v^2$ and $\int ef_e^+(e)\,de = \int ef_e^-(e)\,de$. The shape of the inefficiency distributions under correct and wrong skewness is shown in Panel (a) of Fig. 1, and the corresponding distributions of e in Panel (b). For $\gamma > 0$, the distribution of u (e) has positive (negative) skewness, whereas for $\gamma < 0$ its skewness is negative (positive).

Before deriving the likelihood, it is worth mentioning that two approaches exist to distinguish between correct and wrong skewness. On the one hand, an a priori decision regarding which density to use for restricted maximum likelihood can be made based on the OLS residual. This approach involves estimating the model in (1) by restricting $u_{it} = 0$ through stylized panel data OLS estimation in a first step. Depending on the sign of the OLS residuals, the likelihood can be derived either from (2) if the OLS residuals are negatively skewed, or from (3) if skewness is positive. On the other hand, we follow Hafner et al. (2018) and employ a data-driven approach to selecting γ through one-step unrestricted maximum likelihood estimation. More precisely, the likelihood contains an indicator function to decide between (2) and (3). This approach is more reliable as it avoids the potential biases in skewness that can arise from first-step OLS estimation.

3.1 Marginal ML Estimation

To avoid the necessity of estimating a potentially large number of parameters α_i (for simulation-based evidence of problems related to that, see Greene, 2005), marginal maximum likelihood estimation (MMLE) can be carried out by treating α_i as nuisance parameters and integrating them out. A detailed analysis of asymptotic properties of such estimators is given in Arellano and Bonhomme (2009), see also Bellio and Grassetti (2024) for an application to stochastic frontier models.

We assume in this approach that α_i are random variables. The distribution has to be chosen by the applicant. One possibility would be a uniform distribution over the interval $[-G/2, G/2]$ for some large number G, another one would be a standard normal distribution.

The starting point is the density of the composed error e_{it} that is given in (2) for the case of correct skewness and in (3) for the case of wrong skewness, respectively. Then, the likelihood is given by:

$$\begin{aligned} L(\boldsymbol{\theta} \mid y, \boldsymbol{x}) =& I(\gamma > 0) \prod_{i=1}^{N} \int \prod_{t=1}^{T} f_e^{+}(y_{it} - \alpha_i - \boldsymbol{\beta x}'_{it}; \sigma_v, \gamma) f(\alpha_i)\, d\alpha_i \\ &+ I(\gamma < 0) \prod_{i=1}^{N} \int \prod_{t=1}^{T} f_e^{-}(y_{it} - \alpha_i - \boldsymbol{\beta x}'_{it}; \sigma_v, \gamma) f(\alpha_i)\, d\alpha_i \\ &+ I(\gamma = 0) \prod_{i=1}^{N} \int \prod_{t=1}^{T} f_e^{\cdot}(y_{it} - \alpha_i - \boldsymbol{\beta x}'_{it}; \sigma_v) f(\alpha_i)\, d\alpha_i \end{aligned} \tag{4}$$

where the indicator function $I(\gamma)$ distinguishes between correct and wrong skewness, and also allows for the special case of only fully efficient firms ($\gamma = 0$). In that case, the density of errors $f_e^{\cdot}$ is that of a normal distribution with mean zero and variance σ_v^2. Furthermore, $\boldsymbol{\theta} = (\boldsymbol{\beta}, \sigma_v, \gamma)$ represents a vector of unknown model parameters to be estimated.[3] Following Liu and Pierce (1994), the integral is solved using an accurate adaptive Gauss-Hermite approach, reducing the approximation error for the computation of the integrals to a negligible size (see also Bellio and Grassetti, 2024). Since the distribution of e_{it} is continuous, the integrand function is always smooth. Finally, the likelihood is logarithmized and maximized with respect to $\boldsymbol{\theta}$. As a byproduct of ML estimation, standard errors can be obtained by evaluating the (inverse) Hessian matrix at the maximum.[4]

Because of the indicator function in the likelihood, deriving the asymptotic distribution is not straightforward. However, for either strictly positive or negative γ, the results reported in Arellano and Bonhomme (2009) are valid for independent observations. Arellano and Bonhomme (2009) show that in general the incidental parameter problem is not solved by such integrated likelihood approaches and that there is a finite sample bias for the resulting estimator of $\boldsymbol{\theta}$ of order $\mathcal{O}(T^{-1})$. On the other hand, the bias vanishes if both N and T converge to ∞. As stated by Bellio and Grassetti (2024), however, the estimator based on a marginal likelihood is efficient (subject to some regularity conditions on $\boldsymbol{x}_{it}$). Further, for fixed α_i, the densities f_e^{+} and f_e^{-} fulfill the usual regularity conditions for asymptotic normality, as discussed in Hafner et al. (2018). Intuitively, the conditions still hold if α_i is integrated out as we do here.

[3] Alternatively, the likelihood presented in equation (4) can also be assessed using the extended exponential distribution introduced by Hafner et al. (2018). However, as this study primarily focuses on the half-normal distribution and aims to derive closed-form expressions for the panel case when considering wrong skewness, we abstain from further exploration of the exponential distribution.

[4] The likelihood function is continuous for *fixed* γ, but is not continuous *in* γ in the transition from negative to positive values because of the indicator function. Because all unknown coefficients are simultaneously optimized, which usually requires the likelihood to be continuous, we do not recommend using gradient-based methods but rather the derivative-free simplex method of Nelder and Mead (1965) that is applicable to non-differentiable functions.

3.2 An Approximate Distribution

While the previous approach is straightforward, a potential drawback is the computation time required for solving the integrals. Therefore, we propose a second approach in which the parameters α_i are eliminated directly within the model.

Domınguez-Molina et al. (2003) first recognized the relevance of the CSN distribution for SF modeling. The general form of density of a $CSN_{p,q}(\mu, \Sigma, D, \nu, \Delta)$ distributed p-dimensional random variable Z is $f(z) = \frac{\phi_p(z;\mu,\Sigma)\Phi_q(D(z-\mu);\nu,\Delta)}{\Phi_q(0;\nu,\Delta+D\Sigma D')}$, where ϕ_p is the probability distribution function of the multivariate normal distribution of dimension p, and Φ_q is the cumulative distribution function of the multivariate normal distribution of dimension q. The density in (2) in case of correct skewness is that of a CSN random variable, i.e., $e \sim CSN(0, \sigma^2, -\frac{\gamma}{\sigma_\nu \sigma}, 0, 1)$. However, it remains questionable if the density in (3) in case of wrong skewness can be expressed by parametric distribution as well. Indeed, it turns out that the following approximation, which will be used as an assumption for the subsequent analysis, is useful:

Assumption 1 e^- is closed skew-normal, i.e.

$$e^- \sim CSN_{1\times 2}\left(\underbrace{a_0\gamma}_{\mu_1}, \underbrace{\sigma^2}_{\Sigma_1}, \underbrace{\begin{pmatrix} \gamma/\sigma \\ -\gamma/\sigma \end{pmatrix}}_{D_1}, \underbrace{\begin{pmatrix} -a_0\sigma \\ 0 \end{pmatrix}}_{\nu_1}, \underbrace{\begin{pmatrix} \sigma_\nu^2 & 0 \\ 0 & \sigma_\nu^2 \end{pmatrix}}_{\Delta_1}\right). \tag{5}$$

This assumption implies the density

$$g_e^-(e) = \frac{\phi(e;a_0\gamma, \sigma^2)\Phi_2\left[\begin{pmatrix} \gamma(e-a_0\gamma)/\sigma \\ -\gamma(e-a_0\gamma)/\sigma \end{pmatrix}; \begin{pmatrix} -a_0\sigma \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_\nu^2 & 0 \\ 0 & \sigma_\nu^2 \end{pmatrix}\right]}{\Phi_2\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}; \begin{pmatrix} -a_0\sigma \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma^2 & -\gamma^2 \\ -\gamma^2 & \sigma^2 \end{pmatrix}\right]}. \tag{6}$$

Appendix A discusses from a theoretical perspective, why this assumption is useful, while Appendix B gives numerical evidence for the accuracy of the approximation.

Within each panel i, the distribution of $\boldsymbol{e}_i = (e_{i1}, \dots, e_{iT})'$ is

$$\boldsymbol{e}_i \sim CSN_{T\times 2T}\big(\mu_1 \mathbf{1}_T, I_T\Sigma_1, I_T \otimes D_1, \mathbf{1}_T \otimes \nu_1, I_T \otimes \Delta_1\big), \tag{7}$$

where $\mathbf{1}_T$ is a vector of ones of length T and $\otimes$ is the Kronecker product. This result follows immediately from Corollary 2.4.1 of González-Farıas et al. (2004).

To tackle the incidental parameters problems, we follow the standard econometric literature and use a fixed-effects treatment. Let D denote a fixed-effects transformation matrix to remove the individual-specific effect α_i from the model such as first-differencing (FD) or the within transformation (FE) (Hsiao, 2014). The FD transformation matrix is given by

$$D_{\mathrm{FD}} = \begin{bmatrix} -1 & 1 & 0 & \dots & 0 \\ 0 & -1 & 1 & \dots & 0 \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & 0 & \dots & -1 & 1 \end{bmatrix}_{(T-1)\times T} \tag{8}$$

while the within transformation matrix is:

$$D_{\mathrm{FE}} = \begin{bmatrix} (1-1/T) & (-1/T) & \cdots & (-1/T) \\ (-1/T) & (1-1/T) & \cdots & (-1/T) \\ \vdots & \vdots & \ddots & \vdots \\ (-1/T) & (-1/T) & \cdots & (1-1/T) \end{bmatrix}_{T\times T} \tag{9}$$

Using FD, applied to the model in (1), we obtain

$$D_{\mathrm{FD}}y_{i\cdot} = D_{\mathrm{FD}}\alpha_i + D_{\mathrm{FD}}\boldsymbol{x}'_{i\cdot}\boldsymbol{\beta} + \underbrace{D_{\mathrm{FD}}v_{i\cdot} - D_{\mathrm{FD}}u_{i\cdot}}_{D_{\mathrm{FD}}e_{i\cdot}}, \quad \rightarrow \quad \Delta y_{i\cdot} = \Delta\boldsymbol{x}'_{i\cdot}\boldsymbol{\beta} + \underbrace{\Delta v_{i\cdot} - \Delta u_{i\cdot}}_{\Delta e_{i\cdot}}, \tag{10}$$

where Δ denotes first differences, i.e., $\Delta y_{it} = y_{it} - y_{it-1}$, etc. Using FE, we obtain

$$D_{\mathrm{FE}}y_{i\cdot} = D_{\mathrm{FE}}\alpha_i + D_{\mathrm{FE}}\boldsymbol{x}'_{i\cdot}\boldsymbol{\beta} + \underbrace{D_{\mathrm{FE}}v_{i\cdot} - D_{\mathrm{FE}}u_{i\cdot}}_{D_{\mathrm{FE}}e_{i\cdot}}, \quad \rightarrow \quad \tilde{y}_{i\cdot} = \tilde{\boldsymbol{x}}'_{i\cdot}\boldsymbol{\beta} + \underbrace{\tilde{v}_{i\cdot} - \tilde{u}_{i\cdot}}_{\tilde{e}_{i\cdot}}, \tag{11}$$

with $\tilde{y}_{it} = y_{it} - (1/T)\sum_{t=1}^{T} y_{it}$, etc. Then we can show

Theorem 1 *Under Assumption* 1, *the distributions of* $D_{\mathrm{FD}}e_{it}$ *and* $D_{\mathrm{FE}}e_{it}$ *are CSN*, *i.e.*, $\Delta e_{it} \sim CSN_{(T-1)\times 2T}$ *in* (10) *and* $\tilde{e}_{it} \sim CSN_{T\times 2T}$ *in* (11).

The proof is written down in Appendix C. Then, the corresponding likelihoods are given by:

$$\begin{aligned} L_{FD}(\theta|\Delta\, y, \Delta\, x) =& I(\gamma > 0)\prod_{i=1}^{N} g_e^{+}\left(\Delta y_i - \beta\Delta\, x'_i;\sigma_v,\gamma\right) \\ &+ I(\gamma < 0)\prod_{i=1}^{N} g_e^{-}\left(\Delta y_i - \beta\Delta x'_i;\sigma_v,\gamma\right) + I(\gamma = 0)\prod_{i=1}^{N} f_e\left(\Delta y_i - \beta\Delta x'_i;\sigma_v\right) \end{aligned} \tag{12}$$

$$\begin{aligned} L_{FE}\left(\theta|\tilde{y},\tilde{x}\right) =& I(\gamma;0)\prod_{i=1}^{N} g_e^{+}\left(\tilde{y}_i - \beta\tilde{x}'_i;\sigma_v,\gamma\right) + I(\gamma < 0)\prod_{i=1}^{N} g_e^{-}\left(\tilde{y}_i - \beta\tilde{x}'_i;\sigma_v,\gamma\right) \\ &+ I(\gamma = 0)\prod_{i=1}^{N} f_e\left(\tilde{y}_i - \beta\tilde{x}'_i;\sigma_v\right) \end{aligned} \tag{13}$$

where the indicator function $I(\gamma)$ distinguishes between correct, wrong, and no skewness, i.e., only fully efficient producers. In the latter case, our model collapses to the common fixed-effects estimator. Moreover, $\boldsymbol{\theta} = (\boldsymbol{\beta}, \gamma, \sigma_v)$ is a vector of unknown model parameters to be estimated. The density g_e^+ under correct skewness is given in Chen et al. (2014) and thus our model nests that of Chen et al. (2014). Finally, the likelihood is logarithmized and maximized with respect to $\boldsymbol{\theta}$ (see footnote 4).[5] Note that there is no incidental parameters problem because individual-specific effects α_i do not appear.

As in MMLE, because of the indicator function in the likelihood, deriving the asymptotic distribution is not straightforward. Again, for either positive or negative γ, the estimator should be consistent and asymptotically normal. Note that for $\gamma > 0$, our model collapses to that in Chen et al. (2014), who suppose the estimator features consistency and asymptotic normality. With $\gamma = 0$, our model collapses to the simple fixed-effects estimator, for which asymptotic properties are well known. What is less clear is regarding the case $\gamma < 0$. However, the only difference is that it involves another CSN distribution; g_e^- is still continuous and differentiable for finite σ as g_e^+ is, and the same should be true for the likelihood (again subject to some regularity conditions on $\boldsymbol{x}_{it}$). Accordingly, the estimator should also be consistent and asymptotically normal for $\gamma < 0$.

3.3 Obtaining Efficiency Scores

Computing observation-specific technical efficiency is usually the key goal in empirical application of SF analysis. The conditional expectation estimator suggested by Jondrow et al. (1982) is often adopted for this purpose. Then, the estimation of inefficiency is based on $f(u_{it} \mid e_{it}) = \frac{f(u_{it}, e_{it})}{f(e_{it})}$ evaluated using residuals $\hat{e}_{it}$ of the transformed model, i.e., $\Delta\hat{e}_{it}$ in case of first-differencing or $\hat{\tilde{e}}_{it}$ in case of within transforming (Wang & Ho, 2010), and estimates of $\hat{\gamma}$ and $\hat{\sigma}_v$. Then, the conditional distributions are given by:

$$f^+(u_{it} \mid \hat{e}_{it}) = \frac{f_u^+(u_{it}) f_v(\hat{e}_{it} + u_{it})}{f_e^+(\hat{e}_{it})} \tag{14}$$

$$= \frac{\hat{\sigma}}{\hat{\gamma}\hat{\sigma}_v\sqrt{2\pi}} \exp\left\{ -\frac{\left(u_{it} - \frac{-\hat{e}_{it}\hat{\gamma}^2}{\hat{\sigma}^2}\right)^2}{2\hat{\gamma}^2\hat{\sigma}_v^2} \right\} \Big/ \left(1 - \Phi\left(-\frac{\hat{e}_{it}\hat{\gamma}}{\hat{\sigma}\hat{\sigma}_v}\right)\right) \tag{15}$$

[5] Due to the dependence on the cumulative distribution function (cdf) of a multivariate normal distribution, the CSN densities g_e^+ in equations (12) and (13) necessitate the computation of a $T-1$ and T-dimensional integral, respectively, in case of correct skewness. To simplify these integrals, we utilize the technique described in Appendix C of Chen et al. (2014), taking advantage of the diagonal variance-covariance matrix of the multivariate normal distributions. However, in the case of wrong skewness, where the cdf of the multivariate normal distribution becomes block diagonal, we employ this technique twice.

$$f^{-}(u_{it} \mid \hat{e}_{it}) = \frac{f_u^{-}(u_{it}) f_v(\hat{e}_{it} + u_{it})}{f_e^{-}(\hat{e}_{it})} \tag{16}$$

$$= \frac{\hat{\sigma}}{\hat{\gamma}} \frac{\phi\left(\frac{u_{it} - a_0|\hat{\gamma}|}{|\hat{\gamma}|}\right)}{\Phi(0) - \Phi\left(\frac{-a_0|\hat{\gamma}|}{|\hat{\gamma}|}\right)} \frac{\Phi\left(a_0\right) - \Phi(0)}{\phi\left(\frac{\hat{e}_{it} - a_0|\hat{\gamma}|}{\hat{\sigma}}\right)} \frac{\phi\left(\frac{\hat{e}_{it} + u_{it}}{\hat{\sigma}_v}\right)}{\left[\Phi\left(A_w + \frac{a_0 \hat{\sigma}}{\hat{\sigma}_v}\right) - \Phi\left(A_w\right)\right]} \tag{17}$$

which is evaluated by $\mathbb{E}[u_{it} \mid \hat{e}_{it}] = \int u_{it} f(u_{it} \mid \hat{e}_{it})\, du_{it}$ to get estimates $\hat{u}_{it}$.[6] Following Chen et al. (2014), the integral is solved by means of numerical integration. The sign of the estimated $\hat{\gamma}$ determines whether $f^{+}(u_{it} \mid \hat{e}_{it})$ or $f^{-}(u_{it} \mid \hat{e}_{it})$ is chosen. Finally, efficiency scores obtain as $\widehat{TE}_{it} = \exp\{-\hat{u}_{it}\}$

4 Monte Carlo Simulations

In this section, we report results of Monte Carlo simulations to evaluate the performance of MMLE in comparison with estimation based on the CSN approximation. The scope of this section is twofold. On the one hand, we aim to show that all model parameters can be identified in the panel case, regardless of the sign of error skewness. On the other hand, we assess the validity of the proposed CSN approximation. To emphasize the adverse effects of wrong skewness in panel data analysis, we further compare the proposed estimators with the within-panel estimator by Chen et al. (2014).[7] The data generating process (DGP) follows the related literature about fixed-effects panel SF models (Chen et al., 2014; Belotti & Ilardi, 2018) and reads as:

$$y_{it} = \alpha_i + \beta x_{it} + \epsilon_{it}, \quad \epsilon_{it} = v_{it} - u_{it}, \quad i = 1, \dots, N; t = 1, \dots, T, \tag{18}$$

$$v_{it} \sim \mathrm{N}(0, \sigma_v^2), \tag{19}$$

$$u_{it} \sim \begin{cases} \mathrm{N}_{[0,\infty)}(0, \gamma^2) & \text{if } \gamma > 0, \\ \mathrm{N}_{[0,a_0|\gamma|]}(a_0|\gamma|, \gamma^2) & \text{if } \gamma < 0. \end{cases} \tag{20}$$

[6] As alternatives but less frequently used in practice (Kumbhakar & Lovell, 2003) the conditional distributions can be evaluated at the median or mode; where the latter corresponds to the maximum likelihood estimate for u_{it}.

[7] We also made an attempt to compare the estimator with the 'true fixed-effects model' by Greene (2005), which involves simultaneous estimation of individual-specific effects. However, we encountered significant convergence issues and occasionally obtained peculiar results when applying this estimator in each of the scenarios we consider (which is likely because we consider $N \gg T$). Therefore, we choose not to present the results obtained from this estimator.

$$\lambda = \frac{\gamma}{\sigma_v} \tag{21}$$

In (20), positive (negative) values of γ result in the distribution of u_{it} having positive (negative) skewness. Recall that both distributions have the same expectation. To introduce correlation between individual-specific effects and the explanatory variable, we first generate $\alpha_i, i = 1, \ldots, N$ from a standard normal distribution, i.e., $\alpha_i \sim N(0,1)$ and then obtain x_{it} as $x_{it} = \tau\alpha_i + w_{it}\sqrt{1-\tau^2}$, with $w_{it} \sim \mathrm{N}(0,1)$. Then, x_{it} has mean zero and unit variance; and the correlation between α_i and x_{it} is equal to τ. We set $\beta = 1$, $\tau = .5$, and the variance of ϵ_{it} that is given by $\mathrm{Var}[\epsilon_{it}] = \sigma_v^2 + (\frac{\pi-2}{\pi})\gamma^2$ to one, i.e., $\mathrm{Var}[\epsilon_{it}] = 1$. Distinguished parameter choices for $\lambda = \{-2, -1, 1, 2\}$ induce either positive or negative skewness. In contrast to related studies (Belotti & Ilardi, 2018; Chen et al., 2014; Greene, 2005; Wang & Ho, 2010), we consider combinations of $N = \{10, 20, 50, 100\}$ and $T = \{5, 10\}$ and thus also 'small' sample behavior. The number of replications (R) is 1,000.

The likelihoods given in (4), (12), and (13) are maximized with respect to β, σ_v, and γ. For the estimator based on the CSN approximation, we consider both first-differencing and the within transformation to eliminate the individual-specific effects. Recall that MMLE does not require model transformation because individual-specific effects are integrated out. For the parameter estimates $\hat{\beta}$, $\hat{\sigma}v$, and $\hat{\gamma}$, we report bias, variance, and mean squared error (MSE) statistics. In each replication, we also compute inefficiency estimates according to Jondrow et al. (1982). To calculate MSE, the statistics are averaged first over i for a given replication and then over all replications. An important fact is that all estimators were numerically stable in the simulations and no problems of convergence occurred. This is of particular relevance for estimators based on CSN approximation with large T because numerical optimization of likelihoods based on CSN densities can be cumbersome and might suffer from the 'curse of dimensionality' problem (Belotti & Ilardi, 2018; Chen et al., 2014).[8]

Simulation results including detailed descriptions are given in the Appendix D. In cases of correct skewness, the estimator by Chen et al. (2014) excels in MSE, which is the expected behavior, because there is no statistical need to add uncertainty about the skewness sign. Incidental parameters do not affect slope coefficients. Contrary to expectations, MMLE shows negligible biases in finite T samples. Estimators for β, σ_v, γ and $\mathbb{E}[u|e]$ are consistent across λ, with decreasing MSE as N and T increase. The impact of switching λ is minimal. For wrong skewness, the estimator by Chen et al. (2014) deteriorates rapidly, leading to biased efficiency estimates. This is the expected behavior, as Chen et al. (2014) does not account for wrong skewness. Both newly proposed approaches are consistent, however, and the simulated MSE

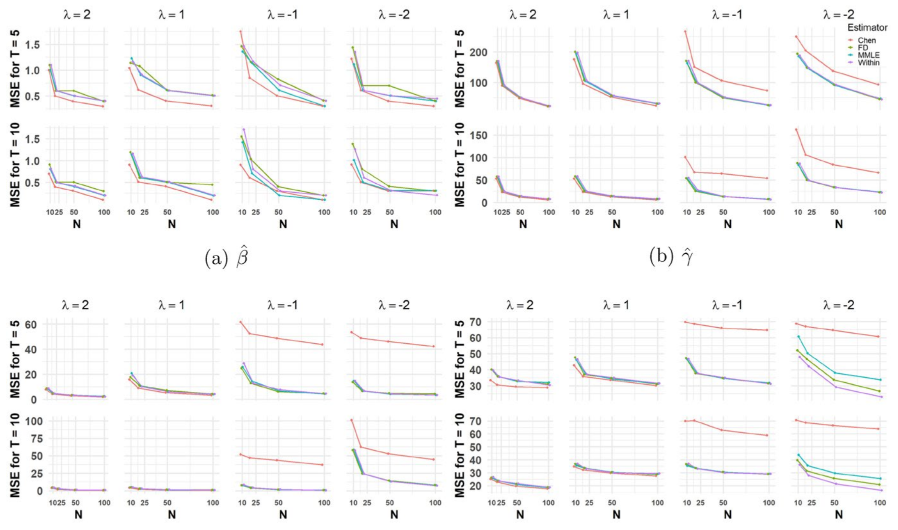
[8] Starting values for the proposed estimators are chosen as follows. For the idiosyncratic variance, the starting value is given by the empirical variance of the dependent variable, i.e., $\sigma v^2 = \widehat{\mathrm{Var}}[y]$; for all other coefficients, we used zeros as starting values. The indicator function therefore takes effect at the starting point via $\mathbb{1}(\gamma = 0)$. This is equivalent to starting with fitting a stylized normal distribution to the dependent variable.

decreases with increasing N or T. MSE differences between MMLE and proposed estimators using CSN approximation are negligible, confirming the proposed method's adequacy in Monte Carlo simulations. Figure 2 additionally shows the simulated MSE in plots for β, σ_v^2, γ and $\mathbb{E}[u|e]$ for varying T, N and λ. While the curves look similar for β across the different estimation methods, the improvement compared to Chen et al. (2014) is clearly visible for σ_v and $\mathbb{E}[u|e]$ in the case of negative λ.

5 Empirical Application

In this section, we illustrate the proposed method to handle (potentially) wrongly skewed inefficiency in panel stochastic frontier models with fixed effects and estimate a production function to quantify the efficiency of the German health care system. SF analysis has become a pertinent approach to analyze the performance of health care systems (for instance, see Evans et al., 2000, Greene, 2004). Greene (2005) motivated his 'true' fixed effects panel SF model by a study undertaken by the World Health Organization and analyzed health care outcomes to unravel efficiencies in the provision of health care services (see also Evans et al., 2000).

We base the empirical application on district-specific health profiles in Germany in the time period from 2004 to 2013. We analyze the population's health by means of an age- and sex-standardized mortality rate (SMR), conditioning on information about the local medical infrastructure, as well as on socio-economic and demographic profiles (for similar applications, see Haschka et al., 2020, Herwartz and



(a) $\hat{\beta}$ (b) $\hat{\gamma}$

(c) $\hat{\sigma}_v^2$ (d) $\mathbb{E}(u|e)$

Fig. 2 MSE statistics. In each panel, the y-axes shows the MSE obtained by the four estimators (red: Chen, green: FD, cyan: MMLE, purple: within) considered subject to the cross-sectional dimension N (x-axes), the time dimension T (row-wisely), and the λ values (column-wisely)

Schley, 2018).[9] Following the related literature (Evans et al., 2000; Greene, 2005; Haschka et al., 2020; Felder & Tauchmann, 2013), we consider each district as a ‘producer’ of health. It is noteworthy, however, that districts do not directly ‘produce’ health in the strict sense of a production process; rather, they provide the legal framework for the structural planning and provision of health care services (Haschka et al., 2020). Following Haschka et al. (2020), the production SFA model for overall health in district i at time t reads as:

$$\begin{aligned}\log(1/SMR)_{it} = \alpha_i &+ \beta_1 \log(gp)_{it} + \beta_2 \log(specialists)_{it} + \beta_3 \log(beds)_{it} \\ &+ \beta_4 \log(dialysis)_{it} + \beta_5 \log(speciality\ diversity)_{it} \\ &+ \beta_6 \log(population\ density)_{it} + \beta_7 \log(gdp\ per\ capita)_{it} \\ &+ \beta_8 \log(unemployment)_{it} + \beta_9 \log(immigrants)_{it} \\ &+ \beta_{10} \log(education)_{it} + v_{it} - u_{it},\end{aligned} \tag{22}$$

$$v_{it} \sim N(0, \sigma_v^2) \tag{23}$$

where $i = 1, \dots, 383$ denotes districts in Germany located in 16 federal states, and $t = 2004, \dots, 2013$ is the year. To capture the medical infrastructure in both the inpatient and outpatient sectors, we consider the number of general practitioners (*gp*), the number of medical specialists (*specialists*), the number of hospital beds per 10,000 inhabitants (*beds*), the number of dialysis devices per 100,000 inhabitants (*dialysis*), and we use the standard deviation within each district across count statistics for five medical specialist groups (internists, ophthalmologists, orthopaedists, psychotherapists per 100,000 inhabitants, and pediatricians per 100,000 children) to generate the variable *speciality diversity*. While the variables *gp* and *specialists* represent the provision of outpatient services, *beds* represent the inpatient sector, *dialysis* accounts for the availability of medical technology, and *speciality diversity* accounts for the heterogeneity of medical service provision within a district. Since past research has shown that socio-economic and demographic factors also influence the population’s health (Felder & Tauchmann, 2013), we further consider the real gross domestic product per capita (*gdp per capita*), the unemployment rate (*unemployment*), the share of employees with a university degree (*education*), and the population share of immigrants (*immigrants*). The first three variables indicate regional deprivation, while the last one pertains to regional diversity. To account for a district’s urbanization, we also take into consideration the *population density*. A more detailed description of the variables used in the above model, including descriptive statistics, is given in Herwartz and Schley (2018), who use the same dataset. However, while they consider log(*gdp per capita*), log(*unemployment*), log(*immigrants*), and log(*education*) as health efficiency determinants (see also Haschka et al.,

[9] As ‘health’ is difficult to quantify, it can only be approximated by, for example, morbidity or mortality measures, for which standardized mortality rate is widely used as a dependent variable in the health literature (for brief reviews, see Balia and Jones, 2008, Bottle et al., 2011).

2020), we directly relate these variables to the health outcome. The panel dataset is (strongly) balanced, and the total number of observations is 3830.[10]

Our specification further diverges from Herwartz and Schley (2018) with regard to the treatment of the inefficiency component u_{it}. Specifically, we allow stochastic inefficiency to vary over i and t and consider both correct and wrong skewness by means of a data-driven choice of distribution of u_{it} in (22), i.e:

$$u_{it} \sim \mathrm{N}_{[0,\infty)}(0,\gamma^2) \quad \text{or} \quad u_{it} \sim \mathrm{N}_{[0,a_0|\gamma|]}(a_0|\gamma|,\gamma^2). \tag{24}$$

The latter case has yet not been considered in empirical health literature and thus offers a novel perspective to unravel structural inefficiencies in the provision of health care services in Germany.

As an alternative specification, one may think of using a truncated normal distribution with positive skewness and (strictly) positive mode, i.e., $u_{it} \sim \mathrm{N}[0,\infty)(\mu,\gamma^2)$ in (22) instead of (24). This means that the majority of health providers do not operate close to full efficiency and thus is in line with our initial considerations that health markets might be characterized by a lack of competition. A flexible choice of the location parameter μ gives a distribution where the majority of providers are either close to full efficiency or further away from it. Although correct skewness is present in either case, this specification would indeed be consistent with our motivation of wrong skewness mentioned in Section 2, according to which it would be found in markets that are not subject to competitive conditions. However, such a specification requires identification of an additional parameter, namely the mode of the distribution (μ). Because simultaneous identification of μ and γ^2 in truncated normal distributions is cumbersome, as both parameters simultaneously determine the expectation and variance of u it (Kumbhakar & Lovell, 2003; Wang & Ho, 2010), the advantage of using (24) is that no additional parameter needs to be identified; the proposed model is therefore parsimonious.

Seeing strong argumentation in favor of state-specific health profiles in Germany (Haschka et al., 2020), we further diverge from Herwartz and Schley (2018) and estimate separate models for each federal state $1,\dots,16$. This strategy allows all parameters, including the skewness of the inefficiency distribution, to be state-specific. Following Kumbhakar and Lovell (2003), we start by assessing the skewness of the OLS residuals prior to the SF analysis to motivate our empirical model that distinguishes between correct and wrong skewness. We thus estimate the above model by restricting $u_{it} = 0$ and use stylized fixed-effects estimation to inspect residual skewness. The spatial distribution of the skewness of the OLS residuals is displayed in Fig. 3. First and unsurprisingly, residual skewness is throughout non-zero, which gives clear evidence in favor of an SF specification. As the sign of residual skewness seems to be state-specific, conducting SF analysis on the federal state level with a heterogeneous treatment of the inefficiency component is presumably justified. For ten out of sixteen federal states in

[10] Note that Germany consists of 401 districts. However, some districts had missing data and needed to be excluded from the analysis, leaving us with 383 districts.

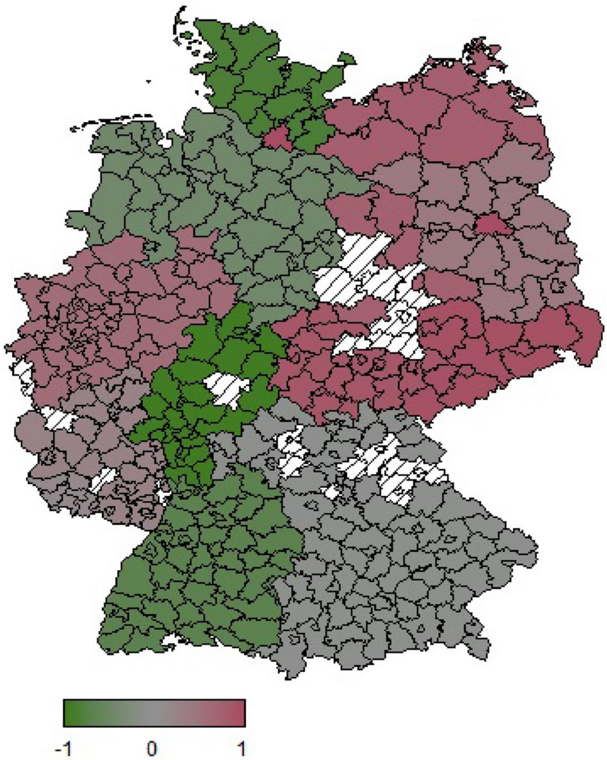


Fig. 3 Skewness of OLS residuals on federal state level. Values are scaled to range between -1 and 1

Germany, the OLS residuals exhibit positive (wrong) skewness, predominantly in West Central and Eastern Germany, while it is negative for the remaining states. This indicates that a classical SF specification that assumes correct skewness is inappropriate to capture the properties of the error distribution in these states.

Seeing substantial evidence in favor of different signs of error skewness, we next estimate the above SF model using MMLE given in (4) and the proposed CSN approximation based on within transformation given in (13), without imposing any sign restrictions on γ. Estimation results are given in Table 1. Unsurprisingly, the effects of health inputs are state-specific; however, some appear counterintuitive. In particular, the effects of variables measuring medical infrastructure sometimes have a negative sign. Although a positive sign is expected because a well-developed medical infrastructure should be beneficial for population health, our findings are in line with related studies (Felder & Tauchmann, 2013; Herwartz & Schley, 2018). For instance, Haschka et al. (2020) explain the negative impact of medical infrastructure by oversupply, which negatively affects population health because people likely receive unnecessary treatments that do not benefit them.

Figure 4 plots the estimated distribution of composite errors and inefficiency. As visible, wrong skewness is evident in West Central and Eastern German states.

Table 1 Parameter estimates for the SFA models for each federal state. Standard errors obtained by the inverse Hessian are given in parentheses

		SH	HH	NI	HB	NW	HE	RP	BW	BY	SL	BE	BB	MV	SN	ST	TH
MMLE	β_1(*General practitioners*)	−.139	−.056	.077	.481	.034	−.146	−.043	−.015	−.021	.036	.048	.086	−.038	−.063	.239	.048
		(.093)	(.084)	(.056)	(.917)	(.062)	(.127)	(.065)	(.064)	(.045)	(.188)	(1.31)	(.101)	(.101)	(.111)	(.133)	(.062)
	β_2 (*Specialists*)	−.026	−.087	.053	.314	.049	.108	−.033	−.099	−.017	.011	.227	.103	−.206	.083	−.091	−.008
		(.072)	(.064)	(.038)	(1.18)	(.051)	(.104)	(.046)	(.056)	(.032)	(.148)	(.286)	(.060)	(.130)	(.054)	(.067)	(.065)
	β_3 (*Hospital beds*)	.035	−.012	−.085	−.785	−.035	−.084	−.011	−.018	−.041	.162	.220	.060	−.110	−.176	−.103	−.008
		(.060)	(.063)	(.028)	(.548)	(.026)	(.044)	(.021)	(.025)	(.020)	(.234)	(1.54)	(.048)	(.095)	(.050)	(.088)	(.031)
	β_4 (*Dialyse*)	.154	.061	.037	−.315	.191	.072	−.039	.266	.119	−.019	.068	−.032	.024	−.049	.029	.047
		(.037)	(.029)	(.0747)	(.391)	(.045)	(.035)	(.030)	(.062)	(.047)	(.097)	(.562)	(.026)	(.021)	(.041)	(.027)	(.041)
	β_5 (*Speciality diversity*)	.014	.051	−.006	−.279	−.003	−.007	.031	−.010	.003	.055	.802	−.039	.060	.102	.025	−.003
		(.033)	(.034)	(.017)	(.497)	(.023)	(.055)	(.022)	(.029)	(.015)	(.062)	(.529)	(.025)	(.051)	(.031)	(.025)	(.025)
	β_6 (*Population density*)	1.10	.687	.530	1.84	1.09	.568	.843	.912	.483	.123	.948	.630	−.194	.666	.680	.567
		(.220)	(.206)	(.132)	(2.43)	(.115)	(.253)	(.162)	(.142)	(.100)	(.703)	(4.12)	(.124)	(.245)	(.123)	(.252)	(.170)
	β_7 (*GDP per capita*)	−.001	.021	.012	.343	.150	−.123	.112	−.002	.014	.026	.497	.023	.120	.041	.016	.023
		(.052)	(.032)	(.035)	(.281)	(.032)	(.078)	(.054)	(.037)	(.028)	(.090)	(.228)	(.042)	(.111)	(.056)	(.107)	(.066)

Table 1 (continued)

		SH	HH	NI	HB	NW	HE	RP	BW	BY	SL	BE	BB	MV	SN	ST	TH
	β_8 (*Unemployment*)	−.066	−.103	−.034	.020	.001	−.014	−.059	.005	−.012	−.083	−.033	−.035	−.058	−.078	−.092	−.037
		(.024)	(.111)	(.018)	(.140)	(.012)	(.025)	(.018)	(.012)	(.010)	(.056)	(.116)	(.029)	(.063)	(.041)	(.050)	(.029)
	β_9 (*Immigrants*)	−.163	−.176	−.015	−.248	−.029	−.091	−.063	−.128	−.010	.060	−.292	−.005	−.011	−.007	−.001	−.014
		(.037)	(.361)	(.017)	(.267)	(.023)	(.056)	(.029)	(.042)	(.018)	(.067)	(.149)	(.018)	(.019)	(.045)	(.043)	(.018)
	β_{10} (*Education*)	.086	.014	.145	.131	.123	.127	.051	.117	.030	1.95E-4	.145	.136	.059	.131	.028	.014
		(.037)	(.116)	(.042)	(.248)	(.028)	(.060)	(.048)	(.032)	(.021)	(.111)	(.062)	(.045)	(.028)	(.047)	(.054)	(.056)
	σ_v	.062	.022	.006	.027	.004	.062	.088	.062	.090	.042	.011	.097	.159	.059	.001	.072
		(1.98E-4)	(.048)	(1.00E-5)	(.610)	(2.01E-4)	(1.90E-4)	(.007)	(3.21E-4)	(.090)	(.001)	(.229)	(.008)	(.220)	(.005)	(2.20E-4)	(.008)
	γ	.077	1.8E-8	.526	-.170	.160	.052	−.039	.074	.062	1.67E-4	−2.21E-6	−.521	.401	−.171	−.374	−.081
		(4.44E−4)	(.008)	(.091)	(.024)	(1.98E-4))	(.011)	(.009)	(.002)	(.016)	(2.77E-4)	(.009)	(.004)	(.004)	(1.08E-4)	(1.92E-4)	(.002)
Proposed CSN	β_1(*General practitioners*)	-.144	−.051	.076	.480	.028	−.150	−.041	−.011	−.018	.040	.051	.085	−.038	−.066	.240	.055
		(.090)	(.081)	(.057)	(.920)	(.061)	(.131)	(.062)	(.061)	(.044)	(.190)	(1.32)	(.106)	(.092)	(.109)	(.134)	(.061)
	β_2 (*Specialists*)	−.021	−.088	.050	.341	.054	.111	−.031	−.104	−.012	.010	.222	.105	−.209	.080	−.096	−.008
		(.070)	(.063)	(.033)	(1.22)	(.050)	(.105)	(.046)	(.055)	(.019)	(.151)	(.290)	(.057)	(.133)	(.052)	(.066)	(.069)
	β_3 (*Hospital beds*)	.031	−.011	−.088	−.780	−.031	−.083	−.017	−.015	−.040	.160	.224	.070	−.114	−.179	−.104	−.009

Table 1 (continued)

	SH	HH	NI	HB	NW	HE	RP	BW	BY	SL	BE	BB	MV	SN	ST	TH
	(.058)	(.062)	(.022)	(.555)	(.020)	(.042)	(.019)	(.019)	(.024)	(.240)	(1.50)	(.050)	(.093)	(.048)	(.090)	(.029)
β_4 (*Dialyse*)	.155	.054	.042	−.291	.251	.070	−.040	.319	.158	−.027	.070	−.030	.021	−.087	.040	.043
	(.033)	(.025)	(.076)	(.405)	(.041)	(.033)	(.024)	(.069)	(.052)	(.110)	(.595)	(.019)	(.029)	(.040)	(.022)	(.037)
β_5 (*Speciality diversity*)	.011	.072	.015	−.381	.019	.012	.024	−.026	1.44E-5	.051	1.02	-.001	.066	.089	.021	−.008
	(.030)	(.028)	(.011)	(.556)	(.029)	(.047)	(.020)	(.018)	(.010)	(.057)	(.556)	(.021)	(.067)	(.025)	(.019)	(.031)
β_6 (*Population density*)	.981	.743	.491	2.09	.891	.480	.666	1.10	.571	.121	.989	.581	−.206	.709	.591	.551
	(.219)	(.200)	(.119)	(2.35)	(.106)	(.220)	(.174)	(.114)	(.109)	(.770)	(3.98)	(.111)	(.229)	(.122)	(.250)	(.161)
β_7 (*GDP per capita*)	.081	.033	.008	.381	.154	−.098	.110	.081	.012	.025	.569	.022	.191	.040	.011	.019
	(.050)	(.029)	(.031)	(.289)	(.038)	(.071)	(.057)	(.033)	(.019)	(.091)	(.218)	(.040)	(.118)	(.048)	(.112)	(.072)
β_8 (*Unemployment*)	−.072	−.083	−.033	.009	.017	.002	−.083	−.002	−.011	−.088	−.018	−.032	−.063	−.075	−.083	−.035
	(.021)	(.105)	(.012)	(.151)	(.010)	(.022)	(.020)	(.015)	(.009)	(.051)	(.110)	(.023)	(.060)	(.037)	(.044)	(.025)
β_9 (*Immigrants*)	−.150	−.202	.004	−.259	−.032	−.103	−.066	−.087	.005	.065	−.290	.004	−.020	−.005	−.001	−.011
	(.033)	(.360)	(.015)	(.261)	(.022)	(.051)	(.031)	(.040)	(.011)	(.064)	(.152)	(.014)	(.020)	(.044)	(.044)	(.016)
β_{10} (*Education*)	.092	.011	.140	.126	.118	.122	.065	.156	.044	−.005	.141	.129	.067	.119	.022	−.002
	(.033)	(.119)	(.040)	(.245)	(.031)	(.062)	(.051)	(.032)	(.020)	(.109)	(.084)	(.041)	(.019)	(.041)	(.051)	(.048)
σ_v	.054	.011	.009	.031	.008	.077	.082	.060	.075	.021	.017	.108	.164	.067	.002	.067

Table 1 (continued)

	SH	HH	NI	HB	NW	HE	RP	BW	BY	SL	BE	BB	MV	SN	ST	TH
	(2.84E−4)	(.052)	(1.92E−4)	(.611)	(1.87E−4)	(1.71E−4)	(.006)	(2.87E−4)	(.099)	(.001)	(.221)	(.008)	(.213)	(.004)	(1.98E−4)	(.007)
γ	.082	1.1E−6	.581	−.092	.151	.049	−.071	.081	.055	1.02E−5	1.86E-7	-.526	.581	-.261	-.400	-.071
	(3.91E-4)	(.009)	(.087)	(.020)	(2.00E-2)	(.008)	(.008)	(1.81E-4)	(.024)	(3.78E-4)	(.012)	(.002)	(.010)	(1.93E-4)	(.002)	(.001)
Number of observations	150	10	460	20	520	250	330	440	890	60	10	180	80	130	80	220

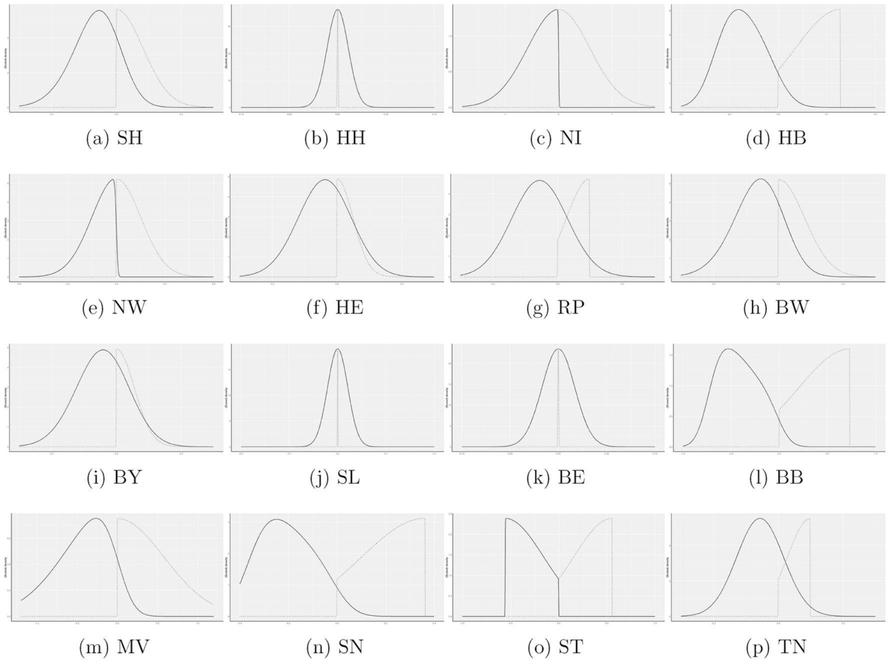
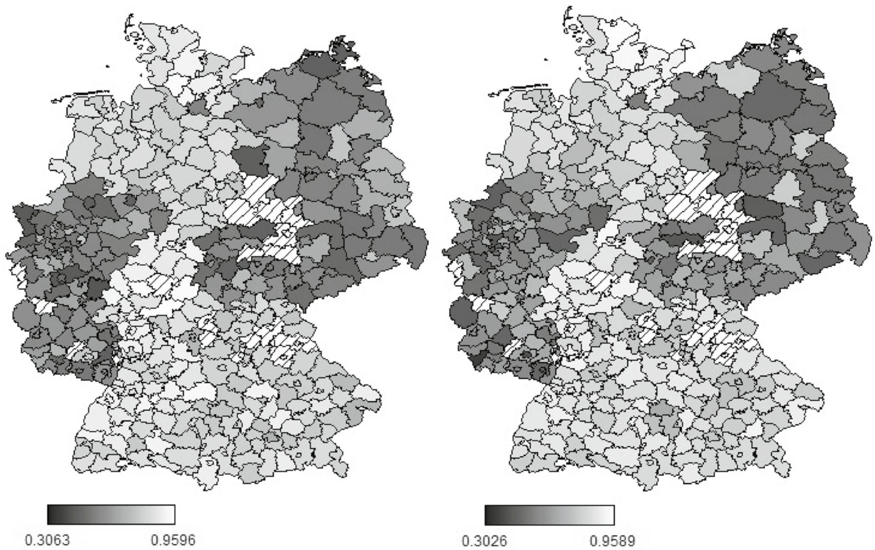


Fig. 4 Estimated density of composite errors (straight lines), and inefficiency (dotted lines) for the federal states. Inefficiency distributions are scaled to have the same peak as the distribution of composite errors

Following that strand of literature that assumes that wrong skewness originates from the inefficiency component (Haschka, 2024b, a; Papadopoulos & Parmeter, 2023), this empirical detection implies that many more inefficient than efficient health care providers are prevalent-something that could not be detected according to the classical SF specification. Our findings likely indicate that, since health care markets are not subject to competitive conditions, the market is unable to solve the problem of wrong skewness on its own. Because the market itself offers no incentives for optimizing the efficiency of health care, it remains up to policy-makers to ensure this.

Estimated technical efficiency scores obtained by the SF models are shown in Fig. 5, respectively for estimation by means of the proposed CSN approximation (Panel (a)) and MMLE (Panel (b)). The predictions of technical efficiency scores show that German districts do not reach their maximum possible level of efficiency. On average, districts reach an efficiency level of 0.704 (min: 0.306; max: 0.960) and 0.722 (min: 0.303; max: 0.957) for the CSN model and the MMLE model, respectively. In line with related studies (Felder & Tauchmann, 2013; Haschka et al., 2020), the efficiency of health care seems high in Southern Germany, Hesse, and North West Germany, while unfavorable conditions prevail in the districts in West Central Germany. By contrast, Herwartz and Schley (2018) and Haschka et al. (2020) report superior efficiency in Eastern Germany. Contrary to their conclusions that the restructuring of medical infrastructure and health provision after the German



(a) Estimation based on proposed CSN approximation

(b) Estimation by means of MML

Fig. 5 Estimated technical efficiency scores obtained by the SF model

reunification is the reason for superior efficiency, our results indicate that wrong skewness might be responsible. Based on our model that considers both wrong and correct skewness, efficiency scores in Eastern Germany are lower in comparison with Western Germany. While the quality of German healthcare is considered generally high (Blümel et al., 2020), there is a consensus in the literature that health care in Eastern Germany still lags behind that in Western Germany (Bambra et al., 2019; Schwartz & Buser, 2005; Brenner, 2001). Possible reasons could be the burdens of the former German Democratic Republic, according to which, even more than thirty years after the German reunification, poor regional health care provision may still be prevalent as a legacy of the planned economy in Eastern Germany. Thus, the result that there are relatively more districts in the West compared to the East that are efficient at providing health care seems plausible.

Previous research has already emphasized that missing incentives likely create vicious cycles of inefficiency (Hill & Miller, 2010; Felder & Tauchmann, 2013). Although the German health care system is among the most consumer-oriented and restriction-free health care systems, it faces rapidly rising costs and serious quality problems (Felder & Tauchmann, 2013). Especially, the regional misallocation of health care infrastructure has led to efficiency losses in health care provision (Ricketts & Holmes, 2007). Given Germany's strong federal structure and the decentralized nature of health policy, special attention should be given to states where wrong skewness is evident, such as West Central and Eastern Germany. Recall that wrong skewness implies that the majority of health care providers are inefficient while only a small number are efficient. In a competitive market, this would not be observed, as competition would incentivize efficiency improvements, resulting in a majority of efficient providers. However, since health care markets lack competitive conditions, they may not be able to address the issue of wrong skewness on their own. Consequently, there is a need for health policy interventions to incentivize efficiency enhancements. Specifically, policy measures should be implemented to promote strategies that reduce unnecessary treatments and encourage evidence-based guidelines, especially in deprived regions like West Central and Eastern Germany. The initiative 'Klug entscheiden' is an important step in this direction (Fölsch et al., 2017). Additionally, incorporating regional socio-economic profiles into the regional planning of medical resource allocation rules would be beneficial (Haschka et al., 2020; Ozegowski & Sundmacher, 2014).

6 Conclusion

Under the traditional SF specification, inefficiency is commonly modeled by a half-normal distribution with positive skewness. However, if true inefficiency is negatively skewed, efficiency scores will effectively be one, leading to false conclusions of high efficiency (Waldman, 1982). Although previous research has emphasized the relevance of the wrong skewness problem in SF analysis (Curtiss et al., 2021; Choi et al., 2021; Wheelock & Wilson, 2020; Daniel et al., 2019), extant studies have only dealt with cross-sectional models. The methodological literature discusses (i) asymmetry of the idiosyncratic noise term, (ii) dependence between idiosyncratic noise and inefficiency, and (iii) wrong skewness of the inefficiency term as potential sources. Against this background, we follow the literature in point (iii); the scope of this paper is to propose a panel data generalization with fixed effects that distinguishes between correct and wrong skewness in one model, without imposing a priori sign restrictions on inefficiency skewness and without having to identify additional parameters that determine skewness.

We start with the cross-sectional SF model by Hafner et al. (2018) and propose an approximation of the distribution of composed errors by means of a CSN distribution if the errors exhibit wrong skewness. Second, we assess the validity of the proposed approximation by inspecting the Kullback–Leibler divergence, which shows that the informational loss is minor and that the approximation is sufficiently accurate. Third, we introduce a panel setting with unobserved heterogeneity and a fixed effects treatment, and we use fixed effects transformations to eliminate the individual-specific effects. We derive the distribution of composed errors after the transformation so that the model can be estimated by means of maximum likelihood. We model inefficiency and heterogeneity separately in the same model to segregate the two effects while allowing for a data-driven choice of positively or negatively skewed inefficiency. Accordingly, our model nests the traditional fixed effects SF model of Chen et al. (2014).

We conduct a set of Monte Carlo simulations on fixed-effects panel SF models and compare the estimator based on the proposed CSN approximation to marginal maximum likelihood estimation (MMLE) based on the exact density, as well as the estimator by Chen et al. (2014). While the latter exhibits severe bias in the case of wrong skewness, the proposed estimators seem to be consistent for all unknown coefficients, including the efficiency score, regardless of whether correctly or wrongly skewed inefficiency is used in the data-generating process. Although MMLE comes with slightly smaller MSE statistics in all scenarios considered, performance differences are minor, which provides additional evidence of the accuracy of the proposed approximation.

Appendix A: Theory of the CSN Approximation

We first make the assumption that the term $\frac{\Phi\left(A_w+\frac{a_0\sigma}{\sigma_v}\right)-\Phi(A_w)}{\Phi(a_0)-\Phi(0)}$ in (3) can be replaced by $\frac{\Phi\left(A_w+\frac{a_0\sigma}{\sigma_v}\right)}{\Phi(a_0)} - \frac{\Phi(A_w)}{\Phi(0)}$. This assumption is useful because it allows deriving a closed-form expression. Then, (3) can be rewritten as:

$$g_e^-(e) = \frac{\phi\left(\frac{e-a_0\gamma}{\sigma}\right)\Phi\left(A_w + \frac{a_0\sigma}{\sigma_v}\right)}{\sigma\Phi\left(a_0\right)} - \frac{\phi\left(\frac{e-a_0\gamma}{\sigma}\right)\Phi\left(A_w\right)}{\sigma\Phi(0)}, \tag{25}$$

which is the difference of two individual $CSN_{1\times1}$ densities; the first component is the density of a $CSN(a_0\gamma, \sigma^2, \gamma/\sigma, -a_0\sigma, \sigma_v^2)$ distribution and the second that of a $CSN(a_0\gamma, \sigma^2, -\gamma/\sigma, 0, \sigma_v^2)$ density. Note that the assumption made to result in (25) does not hold in general and thus only leaves us with an approximation. However, we will show in Appendix B that the difference between (3) and (25) is minor and the approximation is sufficiently accurate.

Using Corollary 4 of Theorem 4 in González-Farıas et al. (2004), and examples of sums of skew-normal random vectors in Genton (2004), another approximation for the partitioned distributions in (25) yields a $CSN_{1\times2}$ distribution (for a similar approach, see Brorsen and Kim, 2013) as described in Assumption 1.

Appendix B: Numerical Accuracy of the CSN-Approximation

We next show that the difference between the 'exact' density f_e^- in (3) and the proposed CSN approximation g_e^- in (6) is minor. For this purpose, we assess the Kullback–Leibler divergence, which is given by

$$\mathrm{D_{KL}} = \int_{-\infty}^{\infty} f_e^-(x) \log \frac{f_e^-(x)}{g_e^-(x)}\, dx. \tag{26}$$

$\mathrm{D_{KL}}$ denotes the expectation of the log difference between the probability distribution with the approximation, i.e., 'how much information do we expect to lose when using the approximation instead of the exact distribution?'. As the shape of the distributions is determined by variance parameters σ_v and γ, we inspect $\mathrm{D_{KL}}(\sigma_v, \gamma)$ subject to different choices for σ_v and γ.

Figure 6 shows $\mathrm{D_{KL}}(\sigma_v, \gamma)$ for various combinations of $\sigma_v \in (0, 8]$ and $\gamma \in [-4, 4]$. What stands out is that the highest values are found around a V-shape with a ratio of

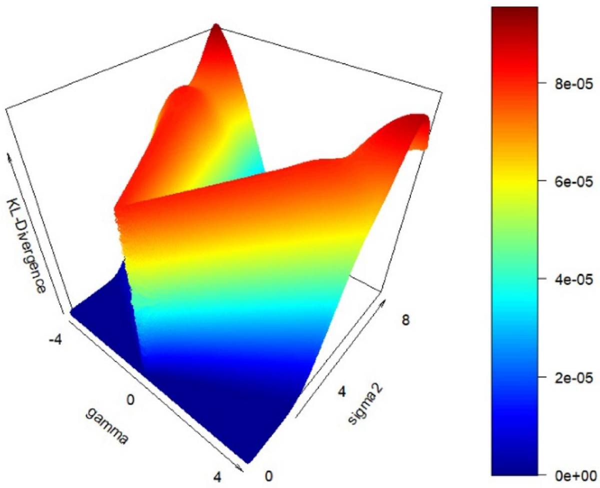


Fig. 6 Kullback–Leibler divergence. $\mathrm{D}_{\mathrm{KL}}(\sigma_v, \gamma)$ quantifies the informational loss from $f_e^-(x;\sigma_v,\gamma)$ to $g_e^-(x;\sigma_v,\gamma)$ depending on σ_v and γ

one to two for γ and σ_v. When $|\gamma/\sigma_v| \rightarrow .5$ the difference between the distributions increases. Nevertheless, the difference is generally small and not greater than about 8E-5, even if $|\gamma/\sigma_v|$ is close to.5. Thus, the informational loss when using g_e^- instead of f_e^- is generally neglectable and we conclude that the proposed CSN approximation is sufficiently accurate to describe the error distribution in case of wrong skewness.

Appendix C: Proof of Theorem 1

Using Proposition 2.3.1 of González-Farıas et al. (2004), linear combinations of jointly CSN random variables are also CSN. Accordingly, if $Z \sim CSN(\mu, \Sigma, P, \nu, \Lambda)$, then

$$AZ \sim CSN(A\mu, A\Sigma A', P\Sigma A'(A\Sigma A')^{-1}, \nu, \Lambda + P\Sigma P' - P\Sigma A'(A\Sigma A')^{-1}A\Sigma P')$$

. With $Z = \boldsymbol{e}_i = (e_{i1}, \ldots, e_{iT})'$ in case of first-differencing and with $A = D_{\mathrm{FD}}$, it is $AZ = \Delta\boldsymbol{e}_i = (\Delta e_{i1}, \ldots, \Delta e_{iT-1})'$. Then, it follows that

$$\Delta\boldsymbol{e}_i \sim CSN_{(T-1)\times 2T}\left(\mu^{\mathrm{FD}}, \Sigma^{\mathrm{FD}}, D^{\mathrm{FD}}, \nu^{\mathrm{FD}}, \Delta^{\mathrm{FD}}\right) \tag{27}$$

$$\mu^{\mathrm{FD}}_{(T-1)\times 1} = A\mu_1 = 0_{T-1} \tag{28}$$

$$\Sigma^{\mathrm{FD}}_{(T-1)\times(T-1)} = \sigma^2 AA' \tag{29}$$

$$D^{\text{FD}}_{2T\times(T-1)} = \left(I_T \otimes \begin{pmatrix} \gamma/\sigma \\ -\gamma/\sigma \end{pmatrix}\right) A'(AA')^{-1} \tag{30}$$

$$v^{\text{FD}}_{2T\times 1} = \mathbf{1}_T \otimes \begin{pmatrix} -a_0\sigma \\ 0 \end{pmatrix} \tag{31}$$

$$\begin{aligned} \Delta^{\text{FD}}_{2T\times 2T} &= \sigma_v^2 I_{2T} + \sigma^2 I_T \otimes \begin{pmatrix} \gamma/\sigma \\ -\gamma/\sigma \end{pmatrix} I_T \otimes \begin{pmatrix} \gamma/\sigma \\ -\gamma/\sigma \end{pmatrix}' \\ &\quad - \sigma^2 I_T \otimes \begin{pmatrix} \gamma/\sigma \\ -\gamma/\sigma \end{pmatrix} A'(AA')^{-1} A I_T \otimes \begin{pmatrix} \gamma/\sigma \\ -\gamma/\sigma \end{pmatrix}' \\ &= \sigma_v^2 I_{2T} + \gamma^2 \frac{1}{T} 1_T 1_T' \otimes \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} \end{aligned} \tag{32}$$

And in case of in case of within transforming, with $A = D_{\text{FE}}$ and $AZ = \tilde{\boldsymbol{e}}_i = (\tilde{e}_{i1}, \ldots, \tilde{e}_{iT})'$, the above results are valid but subject to the change of transformation matrix. Using $A = D_{\text{FE}}$ in (28)-(32), dimension further change to $\mu^{\text{FE}}_{T\times 1}$, $\Sigma^{\text{FE}}_{T\times T}$, $D^{\text{FE}}_{2T\times T}$.

Appendix D: Detailed Monte Carlos Simulation Results

Monte Carlo results are given in Tables 2-3. Table 2 shows the results for correct skewness and Table 3 the ones for wrong skewness. For combinations of $N = 100$, $T = 5, 10$, and $\lambda = 1, 2$, the data-generating process is similar to that in Chen et al. (2014). In these scenarios, the estimator by Chen et al. (2014) is the method of choice as it reveals the smallest MSE statistics among all estimators considered (recall that our model nests that by Chen et al., 2014, which comes at the cost of higher estimation uncertainty). Furthermore, an incidental parameters problem does not affect estimation of slope coefficients for all estimators considered. This result was also found by Chen et al. (2014), Greene (2005), and Wang and Ho (2010) in transformed models. With respect to the integral approximation (MMLE), this is a bit surprising given results by Arellano and Bonhomme (2009), as biases in samples with finite T would be expected. As these cannot be found in our simulations, we conclude that even in small samples biases are generally negligible. For all values of λ, i.e., regardless of having positive or negative skewness, the estimators of β, σ_v, γ and $\mathbb{E}[u|e]$ appear to be consistent as the MSE decreases in N and T. For β and σ_v, the MSE decreases linearly. For γ and in particular $\mathbb{E}[u|e]$, the decay is sometimes more slowly. This might be explained due to the presence of indicator functions which are subject to the parameter γ in the likelihood. Note that also Chen et al. (2014) report rather high and slowly decaying MSE values for estimation of $\mathbb{E}[u|e]$. When switching from $\lambda = 2$ to $\lambda = 1$, MSE statistics remain generally unaffected and small sample biases cannot be detected. In comparison with the estimator by Chen et al. (2014), the proposed approaches reveal slightly higher variance statistics,

Table 2 Monte Carlo results of simulations of the model in (18)–(21) with $\lambda = \{1, 2\}$

				N = 10				N = 20				N = 50				N = 100			
λ	T	Parameter		Chen	MMLE	FD	Within	Chen	MMLE	FD	Within	Chen	MMLE	FD	Within	Chen	MMLE	FD	Within
2	5	β	Bias	.001	.001	.001	.001	.001	−.002	−.002	.001	.001	−.001	.0001	.001	−.001	.001	−.001	−.001
			Var	.010	.010	.011	.012	.005	.007	.007	.006	.004	.005	.006	.005	.004	.004	.005	.004
			Mse	1.00	1.00	1.10	1.10	.500	.601	.600	.600	.400	.502	.600	.500	.300	.402	.400	.403
		σ_v	Bias	.010	.010	.001	.012	.008	.003	−.001	−.005	−.004	.002	.002	−.001	.001	.001	−.001	.002
			Var	.078	.084	.089	.085	.041	.045	.046	.044	.028	.033	.034	.033	.020	.024	.023	.021
			Mse	7.72	8.31	8.81	8.51	4.11	4.40	4.62	4.40	2.80	3.20	3.44	3.20	1.90	2.40	2.30	2.01
		γ	Bias	.005	.007	.006	.006	−.005	.002	.002	−.002	.003	−.005	.003	−.001	−.004	.002	.001	.003
			Var	1.63	1.68	1.70	1.70	.892	.904	.911	.909	.482	.507	.511	.502	.221	.230	.226	.225
			Mse	164	168	170	170	89.5	90.2	90.9	90.9	48.6	50.9	51.0	50.2	22.1	22.9	22.6	22.8
		α	Avg Mse	25.0	25.8	25.5	25.5	21.8	22.2	22.8	22.7	17.5	18.1	18.9	18.4	14.5	15.5	15.9	15.9
		$\mathbb{E}[u\|e]$	Avg Mse	33.5	40.0	40.1	39.7	30.6	35.8	35.8	35.5	29.4	32.7	33.0	33.3	28.9	32.1	31.1	30.4
		time		–	1	.149	.217	–	1	.188	.264	–	1	.388	.376	–	1	.431	.472
	10	β	Bias	.002	−.001	−.004	.003	.005	−.001	.010	−.002	−.003	−.003	.007	.011	−.001	.001	.002	.001
			Var	.007	.009	.009	.009	.004	.005	.005	.004	.003	.005	.005	.005	.001	.002	.003	.003
			Mse	.700	.805	.908	.808	.402	.400	.511	.404	.312	.508	.509	.515	.103	.204	.301	.207
		σ_v	Bias	−.006	.029	−.005	.033	.007	.006	−.015	.012	.004	.001	−.004	−.009	−.003	.001	−.004	−.001
			Var	.039	.044	.048	.047	.019	.026	.030	.029	.009	.010	.016	.013	.005	.009	.011	.009
			Mse	3.85	4.33	4.75	4.60	1.86	2.56	2.86	2.76	.906	1.12	1.35	1.24	.506	.805	.906	.805
		γ	Bias	.010	.016	.009	.020	.006	−.001	.006	−.006	−.005	.004	.001	−.004	−.009	.006	.005	.001
			Var	.533	.567	.578	.580	.239	.247	.250	.251	.124	.137	.140	.141	.066	.078	.082	.079
			Mse	53.0	56.8	57.6	57.6	23.0	24.7	25.0	24.8	12.4	13.8	14.0	14.2	5.99	7.86	7.96	7.75
		α	Avg Mse	17.9	18.5	18.6	18.7	15.4	15.8	16.5	15.8	13.6	14.0	14.1	13.8	12.0	12.7	12.8	12.6
		$\mathbb{E}[u\|e]$	Avg Mse	25.1	25.9	26.2	26.6	22.9	23.7	23.8	23.6	19.7	20.9	21.5	21.6	17.7	18.4	18.9	18.8
		time		–	1	.162	.210	–	1	.190	.254	–	1	.349	.379	–	1	.444	.470

Table 2 (continued)

			N = 10				N = 20				N = 50				N = 100			
5	β	Bias	.004	−.006	.008	−.012	−.003	.004	.0098	−.019	−.004	−.001	.008	−.003	−.005	.001	−.008	.005
		Var	.010	.011	.011	.011	.006	.011	.010	.090	.004	.006	.006	.006	.003	.005	.006	.006
		Mse	1.04	1.23	1.14	1.16	.618	.907	1.08	.910	.406	.607	.612	.606	.308	.506	.515	.505
	σ_v	Bias	.019	.213	.016	.148	−.009	.095	.100	.081	.016	.021	.084	−.001	−.010	.011	.040	.011
		Var	.155	.161	.177	.168	.090	.095	.099	.096	.055	.064	.068	.060	.033	.040	.044	.045
		Mse	15.8	20.8	17.6	18.9	8.87	10.5	11.0	9.97	5.27	6.10	6.99	5.95	3.11	3.95	4.32	4.06
	γ	Bias	.105	.124	.130	.088	.025	.039	.016	.056	.011	.050	.007	.044	−.008	.019	.006	−.028
		Var	1.77	1.93	1.99	1.98	.960	1.11	1.09	1.07	.534	.555	.568	.563	.247	.310	.311	.312
		Mse	176	193	200	197	95.9	105	108	106	53.4	55.8	56.6	56.9	24.4	30.5	31.1	30.9
	α	Avg Mse	25.5	26.8	26.7	25.9	22.3	22.5	23.0	22.8	17.8	18.1	19.2	18.9	14.8	15.5	16.4	15.9
	$\mathbb{E}[u\|e]$	Avg Mse	42.8	46.2	47.6	46.5	35.8	37.6	37.0	37.3	33.6	34.2	34.8	34.9	30.2	31.3	31.0	31.4
	time		–	1	.146	.233	–	1	.191	.276	–	1	.351	.383	–	1	.422	.450

, i.e., correct skewness. The table shows the bias (Bias), variance (Var) of the respective coefficients, and mean squared error (MSE), which is multiplied by 100. 'Chen' denotes the estimator proposed by Chen et al. (2014), 'MMLE' is obtained by maximising the likelihood in (4), and estimates obtained by the proposed CSN approximation are denoted as 'FD' for the maximisation of equation (12), and 'within' for the maximisation of equation (13). The estimates of $\mathbb{E}[u|e]$ are evaluated by respective sample-specific $N \times T$ coefficients $\mathbb{E}[u|e]_{it}(r), i, \dots, N; t, \dots, T$ to obtain $r = 1, \dots, R = 1{,}000$ sample-specific MSE statistics. Estimates of α are subject to the estimates $\hat{\beta}$ and obtain using $\hat{\alpha}_i = \bar{y}_{i.} - \hat{\beta}\bar{x}_{i.}$, where $\bar{y}_{i.}$ and $\bar{x}_{i.}$ denote the respective panel-specific means. Estimates of α are evaluated by respective sample-specific N coefficients $\alpha_i(r), i, \dots, N$ to obtain $r = 1, \dots, R = 1{,}000$ sample-specific MSE statistics. Hence, the row marked values show the average MSE statistic. Calculation time is given relative to MMLE

Table 3 Monte Carlo results of simulations of the model in (18) – (21) with $\lambda = \{-1, -2\}$, i.e., wrong skewness. For further notes, see Table 2

				N = 10				N = 20				N = 50				N = 100			
λ	T	Parameter		Chen	MMLE	FD	Within	Chen	MMLE	FD	Within	Chen	MMLE	FD	Within	Chen	MMLE	FD	Within
-2	5	β	Bias	−.002	−.008	.009	.007	.007	−.004	.001	.001	−.003	.003	.034	.005	.001	−.001	.005	−.022
			Var	.013	.012	.016	.014	.006	.006	.007	.006	.004	.006	.007	.005	.003	.005	.005	.005
			Mse	1.22	1.11	1.44	1.35	.608	.609	.703	.605	.405	.507	.699	.505	.305	.403	.407	.449
		σ_v	Bias	.664	.223	.200	.225	.659	.123	.112	.123	.651	.037	.060	−.004	.628	−.001	.099	.062
			Var	.095	.095	.100	.096	.057	.056	.058	.058	.042	.044	.046	.045	.032	.035	.037	.035
			Mse	53.5	14.7	13.8	14.6	48.9	6.67	6.72	6.78	46.1	4.15	4.77	4.25	42.2	3.35	4.25	3.44
		γ	Bias	.839	.209	.230	.188	.734	.087	.025	.012	.721	−.004	.001	−.056	.672	−.060	.007	−.001
			Var	1.82	1.88	1.89	1.86	1.45	1.45	1.50	1.47	.888	.909	.922	.925	.450	.463	.459	.447
			Mse	250	186	194	188	205	147	149	147	138	90.7	92.4	92.7	93.6	46.7	45.7	44.9
		α	Avg Mse	26.4	26.5	25.9	26.1	23.1	23.4	23.7	23.4	19.7	19.0	19.6	19.2	16.8	16.8	17.1	16.9
		$\mathbb{E}[u\|e]$	Avg Mse	68.8	60.69	52.09	47.96	67.0	50.22	46.77	42.11	64.7	38.17	33.79	29.14	60.6	33.78	26.63	23.08
		time		–	1	.219	.344	–	1	.241	.314	–	1	.278	.310	–	1	.400	.418
	10	β	Bias	−.004	−.001	.007	.008	−.003	.004	.010	.001	.004	−.003	.001	−.005	.001	−.001	.005	−.001
			Var	.009	.011	.019	.013	.005	.005	.008	.007	.003	.003	.004	.003	.004	.003	.004	.003
			Mse	.909	1.01	1.88	1.26	.508	.507	.809	.606	.307	.311	.410	.308	.309	.309	.306	.211
		σ_v	Bias	.674	.079	.100	.013	.677	.029	.050	.001	.644	−.011	−.030	−.001	.622	−.048	.009	−.001
			Var	.557	.599	.574	.577	.211	.245	.248	.241	.120	.136	.144	.132	.067	.077	.085	.074
			Mse	101	6.5	58.2	57.5	62.6	24.5	24.7	24.0	52.9	13.4	14.5	13.0	44.7	7.76	8.15	7.24
		γ	Bias	.889	−.004	.0012	−.021	.765	.033	.046	.011	.714	−.018	.054	−.006	.683	.068	.062	.046
			Var	.833	.834	.876	.868	.480	.499	.500	.496	.344	.337	.340	.335	.208	.229	.229	.227
			Mse	162	83.2	87.6	86.6	106	49.6	49.7	49.4	84.8	33.9	33.9	33.2	66.4	23.0	22.5	22.2
		α	Avg Mse	18.3	18.1	18.9	18.8	15.0	15.5	15.9	15.7	14.3	13.8	13.8	13.4	12.5	12.8	12.8	12.4
		$\mathbb{E}[u\|e]$	Avg Mse	70.6	43.8	39.7	36.1	68.6	35.5	31.5	27.9	66.5	29.6	25.8	21.5	63.9	25.5	20.9	16.4
		time		–	1	.224	.341	–	1	.223	.335	–	1	.275	.337	–	1	.433	.473

Table 3 (continued)

				N = 10				N = 20				N = 50				N = 100			
−1	5	β	Bias	.005	.007	.008	.008	.001	.005	.008	.008	−.006	−.002	−.002	−.001	.003	−.001	−.009	.002
			Var	.018	.036	.044	.036	.007	.016	.023	.016	.007	.009	.009	.008	.004	.004	.005	.005
			Mse	1.55	1.36	1.46	1.46	.805	1.54	1.16	1.27	.507	.611	.825	.708	.308	.306	.411	.408
		σ_v	Bias	.688	.365	.338	.610	.666	.256	.211	.2285	.660	.142	.103	.145	.640	.099	.108	.117
			Var	.156	.127	.132	.130	.088	.083	.086	.080	.059	.051	.055	.059	.033	.036	.038	.036
			Mse	61.7	25.8	24.8	48.7	52.4	14.5	12.8	12.9	48.8	6.76	6.14	7.59	43.8	4.25	4.73	4.55
		γ	Bias	.886	.014	.009	.011	.779	−.001	.009	−.002	.742	−.119	−.011	−.111	.711	.096	.078	.086
			Var	1.89	1.65	1.73	1.74	.900	.998	1.00	.995	.529	.499	.502	.507	.236	.246	.257	.253
			Mse	267	163	171	171	151	99.9	100	99.4	106	51.0	49.8	51.6	73.8	25.4	26.0	25.8
		α	Avg Mse	26.7	26.3	25.6	25.3	22.5	22.1	23.0	23.3	17.5	18.3	19.6	18.7	14.5	14.9	15.7	15.9
		$\mathbb{E}[u\|e]$	Avg Mse	69.8	46.5	47.2	46.8	68.6	37.9	37.8	37.8	66.0	34.6	34.8	35.0	64.7	31.8	31.4	31.2
		time		–	1	.272	.384	–	1	.250	.324	–	1	.270	.313	–	1	.441	.466

which is again most likely because of the indicator function in the likelihoods. The MSE is lower for MMLE compared to the CSN approximation. In the latter case, the within transformation yields lower MSE than the FD transformation because one observation per panel. But, as expected, the computation time is shorter for the CSN approximation.

When examining the case of *wrong* skewness, the findings presented in Table 3 reveal the rapid deterioration of the estimator by Chen et al. (2014). Notably, two key observations emerge from the results. Firstly, the estimates for the regression coefficient β remain unbiased, although there is a slight increase in variances compared to the case of correct skewness. This is because the misspecification of the error distribution is a form of heteroskedasticity, which does not affect the unbiasedness of the slope coefficients. Consequently, the mean squared error (MSE) statistics for α exhibit only minor differences compared to the *correct* skewness scenario, as they rely on the estimates of β. However, significant bias is observed in the estimates for σ_v and γ. Furthermore, increasing the sample size in terms of N or T does not mitigate this bias. Consequently, the estimates of efficiency $\mathbb{E}[u|e]$ are also biased. In contrast, the proposed approaches are generally robust to the presence of both correct or wrong skewness. However, similar to the case of correct skewness, small sample biases arise due to the difficulty of the indicator function in accurately identifying the sign of skewness in such cases. Nevertheless, these biases diminish and the MSE statistics significantly decrease when the sample size N or T is increased. Moreover, the comparison of MSE statistics between the MMLE and the proposed estimators relying on CSN approximation (FD and Within) reveals negligible differences. This suggests that the proposed CSN approximation is adequately accurate, as confirmed by Monte Carlo simulations.

Acknowledgements We are grateful to seminar participants in Bonn and Cologne for helpful comments and suggestions. This article previously circulated under the title “Estimating fixed effects stochastic frontier panel models under ‘wrong’ skewness with an application to health care efficiency in Germany”.

Author Contributions The authors contributed equally to the work.

Funding Open Access funding enabled and organized by Projekt DEAL. The authors declare that no funds, grants, or other support were received during the preparation of this manuscript.

Declarations

Conflict of interest The authors have no relevant financial or non-financial interests to disclose.

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.

Akio, T. (1992). Technical efficiency in japanese industries. In R. e. Caves (Ed.), *Industrial Efficiency in Six Nations*, Cambridge: MIT Press. pp. 31–119.

Almanidis, P., Qian, J., & Sickles, R. C. (2014). Stochastic frontier models with bounded inefficiency. *Festschrift in Honor of Peter Schmidt* (pp. 47–81). Springer.

Almanidis, P., & Sickles, R. C. (2011). The skewness issue in stochastic frontiers models: Fact or fiction? In I. van Keilegom & P. W. Wilson (Eds.), *Exploring Research Frontiers in Contemporary Statistics and Econometrics* (pp. 201–227). Berlin/Heidelberg, DE: Springer.

Arellano, M., & Bonhomme, S. (2009). Robust priors in nonlinear panel data models. *Econometrica, 77*(2), 489–536.

Badunenko, O. & Henderson, D. J. (2023). Production analysis with asymmetric noise. *Journal of Productivity Analysis*, 1–18.

Balia, S., & Jones, A. M. (2008). Mortality, lifestyle and socio-economic status. *Journal of Health Economics, 27*(1), 1–26.

Bambra, C., Smith, K. E., & Pearce, J. (2019). Scaling up: The politics of health and place. *Social Science & Medicine, 232*, 36–42.

Battese, G. E., & Coelli, T. J. (1992). Frontier production functions, technical efficiency and panel data: With application to paddy farmers in India. *Journal of Productivity Analysis, 3*(1), 153–169.

Bellio, R., & Grassetti, L. (2024). Efficient estimation of true fixed-effects stochastic frontier models. *Journal of Productivity Analysis*, 1–8.

Belotti, F., & Ilardi, G. (2012). Consistent estimation of the 'true' fixed-effects stochastic frontier model. *CEIS Working Paper 10*:231

Belotti, F., & Ilardi, G. (2018). Consistent inference in fixed-effects stochastic frontier models. *Journal of Econometrics, 202*(2), 161–177.

Blümel, M., Spranger, A., Achstetter, K., Maresso, A., & Busse, R. et al. (2020). Germany: Health system review.

Bonanno, G., De Giovanni, D., & Domma, F. (2017). The wrong skewness problem: A re-specification of stochastic frontiers. *Journal of Productivity Analysis, 47*(1), 49–64.

Bonanno, G., & Domma, F. (2022). Analytical derivations of new specifications for stochastic frontiers with applications. *Mathematics, 10*(20), 3876.

Bottle, A., Jarman, B., & Aylin, P. (2011). Strengths and weaknesses of hospital standardised mortality ratios. *British Medical Journal, 342*, 116.

Brenner, G. (2001). *Die 'gesundheitsmauer' besteht weiter. Ärzteblatt, 10*, 590–593.

Brorsen, B. W., & Kim, T. (2013). Data aggregation in stochastic frontier models: The closed skew normal distribution. *Journal of Productivity Analysis, 39*(1), 27–34.

Carree, M. A. (2002). Technological inefficiency and the skewness of the error component in stochastic frontier analysis. *Economics Letters, 77*(1), 101–107.

Chen, Y.-Y., Schmidt, P., & Wang, H.-J. (2014). Consistent estimation of the fixed effects stochastic frontier model. *Journal of Econometrics, 181*(2), 65–76.

Choi, K., Kang, H. J., & Kim, C. (2021). Evaluating the efficiency of korean festival tourism and its determinants on efficiency change: Parametric and non-parametric approaches. *Tourism Management, 86*, 104348.

Cohen, B., & Winn, M. I. (2007). Market imperfections, opportunity and sustainable entrepreneurship. *Journal of Business Venturing, 22*(1), 29–49.

Colombi, R., Kumbhakar, S. C., 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.

Curtiss, J., Jelínek, L., Medonos, T., Hruška, M., & Hüttel, S. (2021). Investors' impact on czech farmland prices: A microstructural analysis. *European Review of Agricultural Economics, 48*(1), 97–157.

Daniel, B. C., Hafner, C. M., Simar, L., & Manner, H. (2019). Asymmetries in business cycles and the role of oil prices. *Macroeconomic Dynamics, 23*(4), 1622–1648.

Domınguez-Molina, J. A., González-Farıas, G., & Ramos-Quiroga, R. (2003). Skew-normality in stochastic frontier analysis. *Comunicación Técnica No., I*, 3–18.

El Mehdi, R., & Hafner, C. M. (2024). Panel stochastic frontier analysis with positive skewness. *Computational Economics*, 1–18.

Evans, D. B., Tandon, A., Murray, C. J., Lauer, J. A., et al. (2000). The comparative efficiency of national health systems in producing health: An analysis of 191 countries. *World Health Organization, 29*(29), 1–36.
Felder, S., & Tauchmann, H. (2013). Federal state differentials in the efficiency of health production in Germany: An artifact of spatial dependence? *The European Journal of Health Economics, 14*(1), 21–39.
Filippini, M., & Greene, W. (2016). Persistent and transient productive inefficiency: A maximum simulated likelihood approach. *Journal of Productivity Analysis, 45*, 187–196.
Fölsch, U., Hallek, M., Raupach, T., & Hasenfuß, G. (2017). Resonanz und weiterentwicklung der initiative klug entscheiden. *Der Internist, 6*(58), 527–531.
Genton, M. (2004). *Skew elliplical distributions and their Applications: A Journey beyond normality*. Boca Raton, US: Chapman and Hall/CRC.
González-Farıas, G., Domınguez-Molina, J. A., & Gupta, A. (2004). The closed skew normal distribution. In M. Genton (Ed.), *Skew elliplical distributions and their Applications: A Journey beyond normality* (pp. 25–42). Chapman and Hall/CRC.
Green, A., & Mayes, D. (1991). Technical inefficiency in manufacturing industries. *The Economic Journal, 101*(406), 523–538.
Greene, W. (2004). Distinguishing between heterogeneity and inefficiency: Stochastic frontier analysis of the World Health Organization's panel data on national health care systems. *Health Economics, 13*(10), 959–980.
Greene, W. (2005). Reconsidering heterogeneity in panel data estimators of the stochastic frontier model. *Journal of Econometrics, 126*(2), 269–303.
Hafner, C. M., Manner, H., & Simar, L. (2018). The "wrong skewness'' problem in stochastic frontier models: A new approach. *Econometric Reviews, 37*(4), 380–400.
Haschka, R. E. (2024). Endogeneity in stochastic frontier models with 'wrong' skewness: Copula approach without external instruments. *Statistical Methods & Applications, 33*, 807–826.
Haschka, R. E. (2024). "Wrong'' skewness and endogenous regressors in stochastic frontier models: An instrument-free copula approach with an application to estimate firm efficiency in Vietnam. *Journal of Productivity Analysis, 62*, 71–90.
Haschka, R. E., & Herwartz, H. (2020). Innovation efficiency in European high-tech industries: Evidence from a Bayesian stochastic frontier approach. *Research Policy, 49*, 104054.
Haschka, R. E., & Herwartz, H. (2022). Endogeneity in pharmaceutical knowledge generation: An instrument-free copula approach for Poisson frontier models. *Journal of Economics & Management Strategy, 31*(4), 942–960.
Haschka, R. E., Schley, K., & Herwartz, H. (2020). Provision of health care services and regional diversity in Germany: Insights from a Bayesian health frontier analysis with spatial dependencies. *The European Journal of Health Economics, 21*(1), 55–71.
Herwartz, H., & Schley, K. (2018). Improving health care service provision by adapting to regional diversity: An efficiency analysis for the case of Germany. *Health Policy, 122*(3), 293–300.
Hill, S. C., & Miller, G. E. (2010). Health expenditure estimation and functional form: Applications of the generalized gamma and extended estimating equations models. *Health Economics, 19*(5), 608–627.
Horrace, W. C., Parmeter, C. F., & Wright, I. A. (2023). On asymmetry and quantile estimation of the stochastic frontier model. *Journal of Productivity Analysis, 61*, 1–18.
Hsiao, C. (2014). *Analysis of Panel Data*. Cambridge University Press.
Jondrow, J., Lovell, C. K., Materov, I. S., & Schmidt, P. (1982). On the estimation of technical inefficiency in the stochastic frontier production function model. *Journal of Econometrics, 19*(2–3), 233–238.
Kumbhakar, S. C., Lien, G., & Hardaker, J. B. (2014). Technical efficiency in competing panel data models: A study of Norwegian grain farming. *Journal of Productivity Analysis, 41*, 321–337.
Kumbhakar, S. C., & Lovell, C. K. (2003). *Stochastic Frontier Analysis*. Cambridge University Press.
Kumbhakar, S. C., Parmeter, C. F., & Zelenyuk, V. (2020). Stochastic frontier analysis: Foundations and advances. In R. G. Ray, Subhash C. Chambers and S. C. Kumbhakar (Eds.), *Handbook of Production Economics*. Springer. Chap. 10 – 11.
Liu, Q., & Pierce, D. A. (1994). A note on Gauss Hermite quadrature. *Biometrika, 81*(3), 624–629.
Mester, L. J. (1997). Measuring efficiency at US banks: Accounting for heterogeneity is important. *European Journal of Operational Research, 98*(2), 230–242.
Møllgaard, H. P., & Overgaard, P. B. (2001). Market transparency and competition policy. *Rivista di Politica Economica, 91*(4), 11–64.

Nelder, J. A., & Mead, R. (1965). A simplex method for function minimization. *The Computer Journal, 7*(4), 308–313.
Ortega, M. J. R. (2010). Competitive strategies and firm performance: Technological capabilities' moderating roles. *Journal of Business Research, 63*(12), 1273–1281.
Ozegowski, S., & Sundmacher, L. (2014). Understanding the gap between need and utilization in outpatient care–the effect of supply-side determinants on regional inequities. *Health Policy, 114*(1), 54–63.
Papadopoulos, A., & Parmeter, C. F. (2023). The wrong skewness problem in stochastic frontier analysis: A review. *Journal of Productivity Analysis, 61*, 1–14.
Parmeter, C. F., & Racine, J. S. (2013). Smooth constrained frontier analysis. In X. Chen & N. Swanson (Eds.), *Recent Advances and Future Directions in Causality, Prediction, and Specification Analysis: Essays in Honor of Halbert L* (pp. 463–489). Springer, New York: White Jr.
Redmond, W. (2013). Three modes of competition in the marketplace. *American Journal of Economics and Sociology, 72*(2), 423–446.
Ricketts, T. C., & Holmes, G. M. (2007). Mortality and physician supply: Does region hold the key to the paradox? *Health Services Research, 42*(61), 2233–2251.
Schmidt, P., & Sickles, R. C. (1984). Production frontiers and panel data. *Journal of Business & Economic Statistics, 2*(4), 367–374.
Schwartz, F. W., & Buser, K. (2005). A tale of two Germanys: Health problems in Eastern Germany are clearly related to transition. *British Medical Journal, 331*(7510), 234.
Simar, L., & Wilson, P. W. (2010). Estimation and inference in cross-sectional, stochastic frontier models. *Econometric Reviews, 29*(1), 62–98.
Smith, M. D. (2008). Stochastic frontier models with dependent error components. *The Econometrics Journal, 11*(1), 172–192.
Son, T. V. H., Coelli, T., & Fleming, E. (1993). Analysis of the technical efficiency of state rubber farms in vietnam. *Agricultural Economics, 9*(3), 183–201.
Vargas-Leitón, B., Solís-Guzmán, O., Sáenz-Segura, F., & León-Hidalgo, H. (2015). Technical efficiency in dairy herds from costa rica. *Agronomía Mesoamericana, 26*(1), 01–15.
Waldman, D. M. (1982). A stationary point for the stochastic frontier likelihood. *Journal of Econometrics, 18*(2), 275–279.
Wang, H.-J., & Ho, C.-W. (2010). Estimating fixed-effect panel stochastic frontier models by model transformation. *Journal of Econometrics, 157*(2), 286–296.
Wheelock, D. C., & Wilson, P. W. (2020). In: New estimates of the lerner index of market power for US banks. *Federal Reserve Bank of St. Louis: Working Paper*, 1–32.

Authors and Affiliations

Rouven E. Haschka[1,2] · **Dominik Wied**[3]

✉ Dominik Wied
dwied@uni-koeln.de

Rouven E. Haschka
haschka@zu.de; rouven.haschka@uni-corvinus.hu

1 Chair of Business Analytics & Data Science, Zeppelin University, Am Seemoser Horn 20, D-88045 Friedrichshafen, Germany

2 Institute of Strategy and Management, Corvinus University, Fővám tér 8, H-1093 Budapest, Hungary

3 University of Cologne, Cologne, Germany