



“Wrong” skewness and endogenous regressors in stochastic frontier models: an instrument-free copula approach with an application to estimate firm efficiency in Vietnam

Rouven E. Haschka ^{1,2}

Accepted: 9 March 2024
© The Author(s) 2024

Abstract

Stochastic frontier models commonly assume positively skewed inefficiency. However, if the data speak against this assumption, sample-failure problems are often cited, but less attention is paid to economic reasons. We consider this phenomenon as a signal of distinctive population characteristics stemming from the inefficiency component, emphasizing its potential impact on evaluating market conditions. Specifically, we argue more generally that “wrong” skewness could indicate a lack of competition in the market. Moreover, endogeneity of model regressors presents another challenge, hindering the identification of causal relationships. To tackle these issues, this paper proposes an instrument-free estimation method based on Gaussian copulas to model the dependence between endogenous regressors and composite errors, while accommodating positively or negatively skewed inefficiency through simultaneous identification. Monte Carlo simulation experiments demonstrate the suitability of our estimator, comparing it with alternative methods. The contributions of this study are twofold. On the one hand, we contribute to the literature on stochastic frontier models by providing a comprehensive method for dealing with “wrong” skewness and endogenous regressors simultaneously. On the other hand, our contribution to an economic understanding of “wrong” skewness expands the comprehension of market behaviors and competition levels. Empirical findings on Vietnamese firm efficiency indicate that endogeneity hinders the detection of “wrong” skewness and suggests a lack of competitive market conditions. The latter underscores the importance of policy interventions to incentivize firms in non-competitive markets.

Keywords Stochastic frontier · Wrong skewness · Endogenous regressors · Copula function

1 Introduction

The classical assumption in production stochastic frontier (SF) models is that inefficiency exhibits positive skewness, while the noise term is symmetrically distributed, resulting in the composite error, i.e., the regression residuals, having

negative skewness.¹ However, in empirical applications, the residuals may exhibit positive skewness.² Although the methodological SF literature has proposed some approaches to handle skewness issues (e.g., Hafner et al. 2018), another substantial empirical challenge arises in the case of regressor endogeneity. For example, feedback mechanisms linking output to input can introduce endogeneity if producers adjust their inputs based on the inputs that yield the

✉ Rouven E. Haschka
haschka@zu.de

¹ Chair of Business Analytics and 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

¹ Our focus is on the production SF model. In the cost SF model, the classical assumptions imply positive skewness of the regression residuals.

² Waldman (1982) first demonstrated that if the residuals from the SF model exhibit “wrong” skewness, i.e., positive under the production SF model, inefficiency variance effectively becomes zero. Consequently, efficiency scores tend to one, leading to false conclusions of high efficiency (Hafner et al. 2018; Parmeter and Racine 2013). Green and Mayes (1991) argue that this either indicates “super efficiency” (all firms in the industry operate close to the frontier) or the inappropriateness of the SF analysis technique to measure inefficiencies.

highest marginal outcome (Siebert 2017). Endogeneity can also be introduced if firms know their own inefficiency and adjust their inputs accordingly, but this is unobserved to the analyst (Haschka and Herwartz 2022). While there are many economic reasons for endogeneity, potential economic understandings of “wrong” skewness are far less explored.

The occurrence of “wrong” skewness has been attributed to poor samples (Almanidis and Sickles 2011; Hafner et al. 2018). Simar and Wilson (2009) confirm that small samples can lead to incorrect skewness measures despite correct skewness in the underlying population. To address this issue, researchers should prioritize increasing the sample size rather than making changes to the model specification (Almanidis and Sickles 2011, p. 201). However, what if “wrong” skewness occurs in large samples, such that explaining it by poor sampling or bad luck might not be justified? Methodologically, three reasons for “wrong” skewness have been considered in the literature: (i) asymmetry of idiosyncratic noise, (ii) dependence between idiosyncratic noise and inefficiency, or (iii) “wrong” skewness of the inefficiency component. In contrast, existing economic explanatory approaches for “wrong” skewness deal with the issue in general (for a recent summary, see Papadopoulos and Parmeter 2023), but do not delve into investigating where (i.e., from which component) it might originate.

By assuming that “wrong” skewness is due to the inefficiency component, our contribution to the economic explanation aims to question the characteristics of the market. If “wrong” skewness is detected in a market that is expected to be competitive, we argue that it suggests that market competition is not generating adequate incentives for producers to improve their efficiency, thus indicating that the market may not be as competitive as expected. In the absence of competitive pressure that forces producers to increase efficiency to avoid falling behind competitors, we might observe many inefficient producers and only a few efficient ones. In such situations, assuming “wrong” skewness is due to the inefficiency term offers explanations for competition levels and market dynamics.

Since any empirical detection of “wrong” skewness requires estimation, this depends on the performance and consistency of the estimation method used. Even if the underlying (production) model is correctly specified, endogeneity can lead to biased results, making it difficult to detect “wrong” skewness. Endogeneity can be loosely defined as regressor-error dependence, which is particularly important for SF models, as this dependence can stem from correlation with inefficiency or idiosyncratic noise (Griffiths and Hajargasht 2016; Mutter et al. 2013). While the linkage of endogenous covariate information and composite errors can not only lead to biased estimates for causal effects if applied methods build upon assumptions of regressor exogeneity, but the residual distribution may also be distorted,

making it difficult to capture skewness correctly. The standard approach to handle the endogeneity problem in SF models is to use likelihood-based instrumental variable (ML-IV) estimation methods (Amsler et al. 2016; Haschka and Herwartz 2022; Kutlu 2010; Prokhorov et al. 2021; Tran and Tsionas 2013). However, a general drawback of ML-IV methods is that they rely on the availability of consensual outside information to construct the instruments.

Against this background, we propose an estimation approach to handle endogenous regressors while simultaneously identifying “correct” or “wrong” skewness to assess market competition levels as an “empirical test of concept”, aiming to contribute to the *economic* understanding of “wrong” skewness. Given the often limited availability or weakness of suitable instrumental variables (IVs) in SF models, we conceptualize our approach as a methodological extension of the IV-free joint regression model using copulas introduced by Park and Gupta (2012). The core idea is to construct the joint distribution of the endogenous regressor and composite error, enabling simultaneous identification of “correct” or “wrong” skewness without the need for IVs. While copula-based endogeneity corrections have been extensively studied and successfully applied in classical SF models (Karakaplan and Kutlu 2015; Tran and Tsionas 2015; Tsionas 2017), their general applicability in SF models with “wrong” skewness remains unexplored. Our proposed approach builds upon and generalizes several existing methods, integrating models presented by Hafner et al. (2018), Park and Gupta (2012), and Tran and Tsionas (2015) into a unified framework. We conduct a series of Monte Carlo simulation experiments to demonstrate the suitability of the proposed estimator and compare it with alternative methods. Since there is currently no method capable of simultaneously addressing both endogeneity and skewness issues, our comparison includes methods that assume exogeneity (Hafner et al. 2018) or “correct” skewness (Tran and Tsionas 2013, 2015).

The empirical application aims to provide an unbiased understanding of the determinants of firm performance using data from 16,474 Vietnamese firms in 2015. Our findings lead to three major conclusions. First, we identify significant regressor endogeneity, challenging the estimation of firm productivity in Vietnam. Under the exogeneity assumption, marginal effects are overestimated, suggesting increasing returns to scale (RTS). However, accounting for endogeneity reveals constant RTS, aligning with the Vietnamese government’s priority of steady growth rates over rapid expansion. Second, the detection of “wrong” skewness is hindered by endogeneity, as explored within the Monte Carlo simulations. Despite lower-than-implied efficiency levels, accounting for endogeneity without addressing “wrong” skewness results in even lower efficiency levels, as the skewness might be falsely attributed to endogeneity. Third,

empirical evidence points to moderate efficiency levels and “wrong” skewness, indicating a growing number of inefficient firms in the market, which contradicts the assumption of competitiveness, given our considerations that “wrong” skewness is due to market forces are valid. This lack of incentives to improve efficiency may be attributed to factors such as corruption and the constraints of the communist regime, hindering the establishment of liberal and competitive market conditions. Policy interventions are therefore necessary to create incentives for firms to optimize their processes and enhance efficiency.

The paper begins by discussing the presence of “wrong” skewness in competitive markets, followed by a brief review of the (IV-free) SF literature addressing endogeneity. Section 3 introduces the model and discusses the copula approach to handle regressor endogeneity in SF models with “wrong” skewness when instrumental information is unavailable. In Section 4, we assess the finite sample performance of the proposed approach through Monte Carlo simulations. The empirical application is detailed in Section 5, followed by the concluding remarks in Section 6.

2 Background

In this section, we first summarize potential explanations for the occurrence of “wrong” skewness which have been discussed in the literature. Subsequently, we elaborate on economic perspectives that are based on the informative nature of detecting “wrong” skewness in (competitive) markets. Finally, we provide a literature overview focusing on endogeneity in SF models, with particular emphasis on instrument-free approaches.

2.1 Reasons for “wrong” skewness

Since “wrong” skewness has primarily been considered an empirical phenomenon (Almanidis and Sickles 2011; Hafner et al. 2018; Waldman 1982), the prevailing reason for its detection is often attributed to small sample sizes (Simar and Wilson 2009, pp. 8–9). In cases where the true skewness is correct but “wrong” skewness is observed in a small sample, an inadequate sample size is typically identified as the cause. However, other reasons behind detecting “wrong” skewness have received less attention (for an excellent recent review, see Papadopoulos and Parmeter 2023). While there are some studies that detect “wrong” skewness in empirical applications (e.g., Almanidis and Sickles 2011; Hafner et al. 2018; Parmeter and Racine 2013), they do not delve into discussing potential characteristics in the population for this finding (one exception is Haschka and Wied 2022).

On the one hand, certain characteristics of the data structure may contribute to “wrong” skewness. The asymmetry of

the idiosyncratic error term can lead to multimodality in the distribution of efficiency scores, which in turn can cause “wrong” skewness (e.g., Badunenko and Henderson 2024; Bonanno et al. 2017; Horrace et al. 2024; Son et al. 1993). Additionally, unmodeled dependence between idiosyncratic noise and inefficiency can be another contributing factor (e.g., Bonanno et al. 2017; Bonanno and Domma 2022; Smith 2008). On the other hand, from an economic perspective, specific characteristics of the underlying population, such as unique features of the market in which firms operate, could also explain “wrong” skewness. As suggested by Papadopoulos and Parmeter (2023), when encountering skewness issues, researchers are advised to first consider the market’s specific attributes and potential peculiarities. Subsequently, they should reassess their arguments and determine whether the skewed result reflects an inherent characteristic of the population or is merely a consequence of a flawed sample.

In reviewing these contributions, two points stand out. First, the literature that discusses methodological reasons for the occurrence of “wrong” skewness fails to relate them to characteristics in the population. Specifically, explaining market mechanisms that introduce dependence between idiosyncratic noise and inefficiency or cause asymmetry in the distribution of idiosyncratic noise requires a sound economic understanding. For instance, to what extent should unobserved production shocks simultaneously increase (or reduce) efficiency, and what accounts for the prevalence of positive shocks over negative ones (or vice versa)? Second, if we start addressing skewness issues more generally by discussing economic reasons, the question arises as to which of the model components requires an adjustment to reflect this peculiarity. Beyond the dependence within the composite error or asymmetry of idiosyncratic noise, skewness issues can also be attributed to inefficiency. What insights can we derive regarding market characteristics by presuming that “wrong” skewness stems from the inefficiency component?

2.2 Economic explanations and market competition

Stochastic frontier models often assume positive skewness in the inefficiency distribution to align with competitive market dynamics in which producers operate (Aigner et al. 1977). This assumption is grounded in economic reasoning, particularly in microeconomic production models that assume producers strive to optimize output given the inputs they use. In competitive markets, producers should minimize costs and maximize outputs, thus operating near the efficiency frontier due to competitive pressure that incentivizes efficiency improvements. Inefficiency is seen as deviations from the frontier, with highly inefficient producers likely exiting the market (Haschka and Herwartz 2020). Thus, specifying stochastic frontier models with positively skewed inefficiency distributions, such as the common half-normal distribution

(Kumbhakar et al. 2020), is justified for evaluating producer efficiency in competitive markets.

What if only a small fraction of the firms attain a level of productivity close to the frontier while a large fraction attains considerable inefficiencies? According to Carree (2002), such a situation that is at odds with the assumption of positively skewed inefficiency might be found in industries characterized by alternating cycles of innovation and imitation, with periods in which a few firms innovate and improve their efficiency, while many firms remain inefficient, yielding “wrong” skewness. In subsequent periods, these firms imitate the innovations and efficiency levels converge, yielding correct skewness. However, these examples impose specific requirements on the market, such as the necessity for innovations to drastically and suddenly enhance efficiencies, the occurrence of these “leapfrog innovations” in a cyclical manner, and that innovation markets are distinguished by a clear distinction between innovation leaders and followers.

Furthermore, Torii (1992) mentions that “wrong” skewness in inefficiency results from technological progress and the non-immediate replacement of assets within each producer, leading to misalignment when a few firms quickly renew their capital stock while the majority do so slowly. However, this implies that “wrong” skewness disappears in the long run once all firms have renewed their capital stock (Torii 1992). This suggests that increased competitive pressure leads to a more rapid renewal of capital stock by firms, resulting in a shorter duration for the phenomenon of “wrong” skewness to be observable.

Both of these explanations implicitly assume that the inefficiency term is responsible for the skewness issues, albeit without explicitly labeling it as such. While they describe very specific situations, our explanatory approach more generally aims to outline that the existence of negative skewness in the inefficiency distribution contradicts the expectations of a competitive market environment. Given the absence of other reasons for “wrong” skewness (e.g., poor samples), an observation that the majority of producers operate at lower efficiency levels, with only a few operating close to the frontier, could be explained by limited competition in the market, where producers may not face sufficient pressure to operate at their maximum efficiency levels (Haschka and Herwartz 2022). Papadopoulos and Parmeter (2023) mention markets with heavy regulation or entry barriers as potential reasons why the majority of established firms sit comfortably near higher inefficiency values without seeing a need to reduce inefficiency.³ Moreover, factors such as limited market transparency (Møllgaard and Overgaard 2001), technological constraints (Ortega 2010), market imperfections (Cohen and

Winn 2007), seller’s markets (Redmond 2013), or structural reasons (Haschka and Wied 2022) could contribute to this lack of competition. Unlike oligopolistic markets, the absence of competition does not necessarily result from market concentration or a small number of producers. In the absence of competitive market mechanisms, producers lack the necessary incentives to improve their efficiency levels.

To illustrate these considerations and our contribution to the economic reasoning that explains “wrong” skewness by attributing it to the inefficiency component, Fig. 1 depicts (potential) reasons and implications that emerge from skewness issues discussed in the literature. While no concerns are expressed if the skewness is correct,⁴ methodological, economic, and small sample sizes have been identified as reasons for “wrong” skewness in the literature. While economic explanatory approaches deal with the problem in general and do not discuss which model component could be responsible, methodological explanatory approaches do not inquire about the economic causes. Likewise, attributing it solely to poor samples or data issues is an oversimplification, especially if it appears in larger samples (in small samples, however, it can never be ruled out with acceptable degree of certainty that the sample is poor). In contrast, we suggest a lack of competitive pressure and associated incentives for producers to enhance their efficiency levels as another potential explanation. Since our explanation is aimed at “wrong” skewness originating from the inefficiency term, existing economic explanations in the literature can be linked (Carree 2002; Papadopoulos and Parmeter 2023; Torii 1992). Assuming the inefficiency term is the source, empirical identification of negative skewness likely provides valuable insights into the general competitive market dynamics (indicated by the green blocks in Fig. 1). A more detailed examination should then follow to determine why there are no incentives to increase efficiency in this market.

2.3 Endogeneity and the use of copulas in SF models

While skewness issues require careful model specification when estimating SF models, endogeneity can greatly hinder the detection of causal effects.⁵ Regressor endogeneity can have various causes. If producers have some a priori information on potentially inefficient output generation, it seems likely that the choice of production inputs is adjusted

³ While this explanation could also apply to the scenario described by Torii (1992), in the example by Papadopoulos and Parmeter (2023), “wrong” skewness may also persist in the long run.

⁴ As shown in the Monte Carlo simulations in Section 4, endogeneity can lead to correct skewness being detected in a sample even though the true skewness is “wrong”. This is relevant because “wrong” skewness can also be justified in the population (Haschka and Wied 2022).

⁵ For recent approaches to cope with endogeneity in nonlinear models (including SFA), the reader may consult the special volume of the *J Econom.* entitled *Endogeneity Problems in Econometrics*, edited by Kumbhakar and Schmidt (2016).

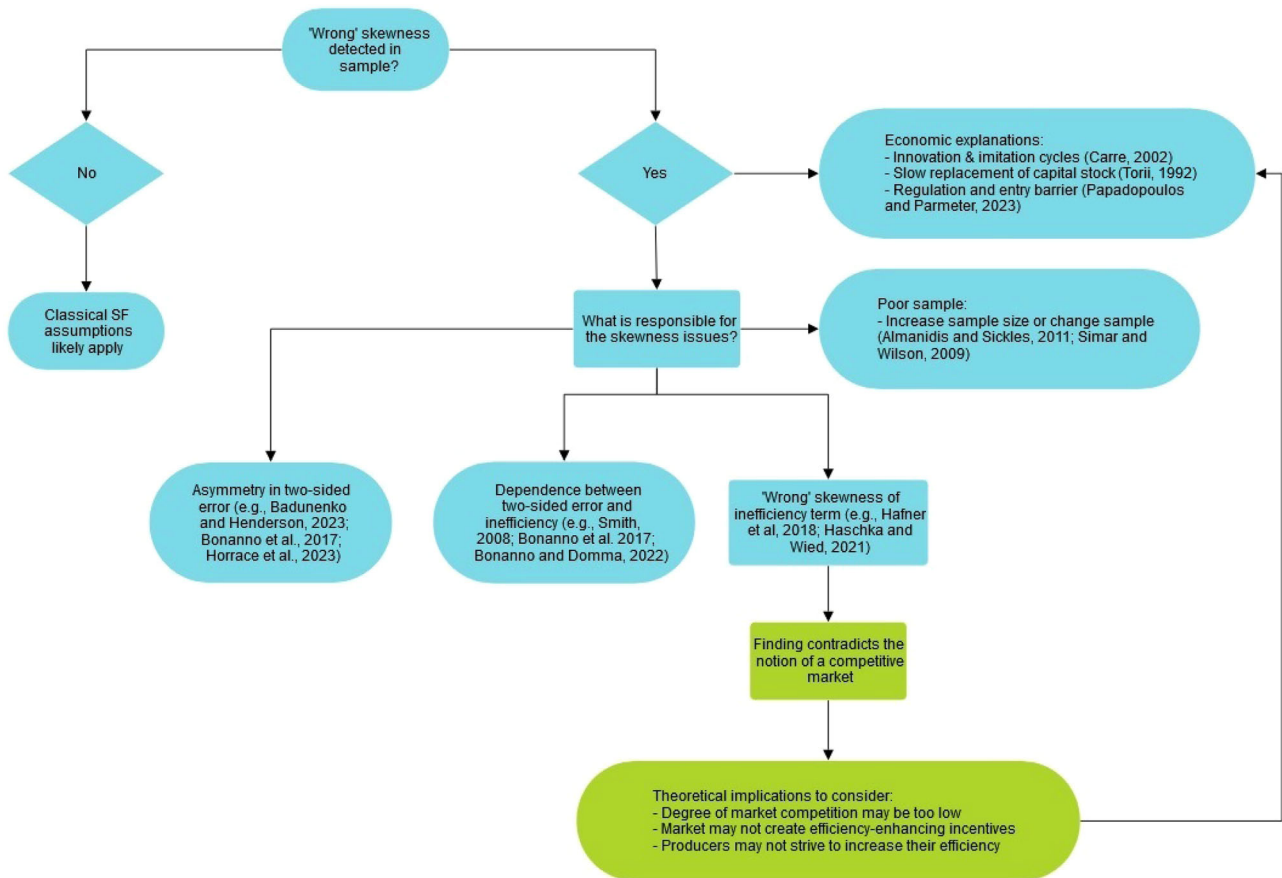


Fig. 1 Implications of detecting correct and “wrong” skewness that are discussed in the literature. The literature references include only studies that refer to “wrong” skewness. What we derive from an empirical detection of “wrong” skewness is shown in green

accordingly (Haschka and Herwartz 2022). In effect, the described unobserved correlation between production input factors and stochastic inefficiency is among the most common forms of endogeneity in the context of efficiency modeling (Cincera 1997). More generally, since output generation is typically seen as a reflection of input activities, successful output generation might also lead to further input activities, inducing endogeneity as a result of patterns of reverse causality. Furthermore, technology shocks that affect investment decisions provide a third origin of endogeneity. Since such shocks are unobserved to analysts, they manifest in model terms assessing productive efficiency (Haschka and Herwartz 2020). Endogeneity bias might also occur when firms respond to demand or supply shocks (that are unobserved to the analyst) by adjusting their inputs, such as the number of employees (Ehrenfried and Holzner 2019). For instance, global health shocks, energy crises, or political tensions might trigger unexpected hiring or investment decisions (Reeb et al. 2012). Lastly, the presence of omitted variables, such as subsidies large enough to have a significant impact on output generation, can also give rise to endogeneity bias.

Traditional approaches for dealing with endogenous regressors in stochastic frontier settings often involve

instrumental variable estimation (Amsler et al. 2016; Griffiths and Hajargasht 2016). These methods utilize exogenous instruments to exploit their informational content and typically employ two-stage-least squares (Amsler et al. 2016; Griffiths and Hajargasht 2016), ML-IV (Amsler et al. 2016; Haschka and Herwartz 2022), control functions (Centorrino and Pérez-Urdiales 2023), or GMM (Shee and Stefanou 2015; Tran and Tsionas 2013) for estimation. However, the validity of instruments remains debatable. Instruments may be scarce, weak, or even unavailable, prompting researchers to explore IV-free alternatives.

The use of copulas has gained increasing attention in stochastic frontier settings, although many applications are not aimed at endogeneity corrections. Amsler and Schmidt (2021) identify three different motivations for the use of copulas in the SF literature: (i) allowing idiosyncratic noise and inefficiency to be correlated in an otherwise standard SF model (e.g., Amsler et al. 2016, 2017; El Mehdi and Hafner 2014; Smith 2008; Wiboonpongse et al. 2015);⁶ (ii) allowing dependence between different composite errors

⁶ This allows addressing one of the potential reasons for “wrong” skewness (see Fig. 1).

and/or other types of errors; for example, to model autocorrelation in panel data (e.g., Amsler et al. 2014; Das 2015; Lai and Kumbhakar 2020), or across different equations in a multi-equation model (e.g., Carta and Steel 2012; Haschka and Herwartz 2022; Huang et al. 2018); and (iii) allowing non-standard types of dependence between the errors in a multi-equation system (e.g., Amsler et al. 2021).

The potentials arising from (i) and (ii) have led to the possibility of taking the endogeneity of regressors into account. That is, it allows for correlation between the regressors and idiosyncratic noise and/or inefficiency (Amsler et al. 2016, 2017). Following instrumental variable theory, Amsler et al. (2016) assume that the endogenous regressor can be decomposed into a part that is correlated with the error and a part that is truly exogenous (i.e., the instrument). Because this decomposition introduces a new equation, this class of models may be seen as multi-equation-type. They use the Gaussian copula to obtain the joint distribution of this correlated part, idiosyncratic noise, and the inefficiency term. Amsler et al. (2017) generalize this approach and allow for environmental variables to affect inefficiency.

To avoid the assumption of decomposability of the endogenous regressors and therefore not belong to the class of multi-equation models, copula approaches directly model regressor-error dependence, and are increasingly explored in SF settings. This class of models uses copula functions to approximate the joint distribution of endogenous regressors and composite errors without requiring instruments. In the first step, data-driven cumulative distribution functions (cdfs) of endogenous regressors are obtained. These, along with an assumed distribution for composite errors, are used as plug-in estimates for the copula function in the second step, and estimates are derived based on the joint distribution. Tran and Tsionas (2015) directly construct this joint distribution using empirical cdfs and Gaussian copula (see also Tsionas 2017). Karakaplan and Kutlu (2015) rearrange the model proposed by Tran and Tsionas (2015) and show that targeting the joint distribution is not necessary because, with two-stage generated regressors, focusing on the marginal distribution of composite errors for ML estimation is sufficient. Papadopoulos (2021) develops a two-tier SF model to handle latent variables, building on the copula approach of Tran and Tsionas (2015).

Note that in the first strand of literature, copulas are employed instead of a closed-form expression for the likelihood function to model the dependence between the error terms, yet instruments are still utilized. This distinction is crucial because this literature avoids many of the identification difficulties encountered by the second stream of literature. The identification problem the second strand faces arises if endogenous regressors have the same distribution as the errors, or if the distributions are very close. In that

case, model identification without IVs breaks down because copulas fail to distinguish noise from variation due to endogenous regressors (Tran and Tsionas 2015).

Assuming normality of the idiosyncratic noise component and the half-normal distribution for inefficiency is a natural choice, since these assumptions lead to the closed skew normal distribution (CSN) for composite errors (Dominguez-Molina et al. 2003; González-Farías et al. 2004); a distribution that is well-defined parametrically. Since joint estimation using copulas requires both the cumulative distribution function (cdf) and the probability density function (pdf) of the composite error distribution, the use of the CSN distribution allows the approach to be implemented in a straightforward manner.

The fact that little attention is paid to skewness issues in SF models becomes even clearer when reviewing these contributions, since extant studies entirely base endogeneity-robust SF modeling on the assumption of “correctly” skewed inefficiency. To our knowledge, regressor endogeneity and “wrong” skewness have not been simultaneously addressed so far. Therefore, we aim to offer a simple solution to this issue by building on previous work by Hafner et al. (2018), Park and Gupta (2012), and Tran and Tsionas (2015).

3 Copula-based handling of endogenous regressors under “wrong” skewness

Consider the typical stochastic frontier model:

$$y_i = x_i' \beta + z_i' \delta + \underbrace{v_i - u_i}_{e_i}, \quad i = 1, \dots, n, \quad (1)$$

where y_i is the output of producer i , x_i is $L \times 1$ vector of exogenous inputs, z_i is $K \times 1$ vector of endogenous inputs, β and δ are $L \times 1$ and $K \times 1$ vector of unknown parameters, respectively, v_i is a symmetric random error, u_i is the one-sided random disturbance representing technical efficiency, and the composite error is therefore $e_i = v_i - u_i$. We assume that x_i is uncorrelated with v_i and u_i , but z_i is allowed to be correlated with v_i and possibly with u_i , and this generates the endogeneity problem. We also assume that u_i and v_i are independent and leave the skewness of u_i unrestricted. The discussion that follows can be easily extended for the case where the (exogenous) environmental variables are included in the distribution of u_i (Battese and Coelli 1995; Haschka and Herwartz 2022).

3.1 “Wrong” skewness of inefficiency distribution

According to standard SF practices, we assume that $v_i \sim N(0, \sigma_v^2)$ captures two-sided idiosyncratic noise. Following

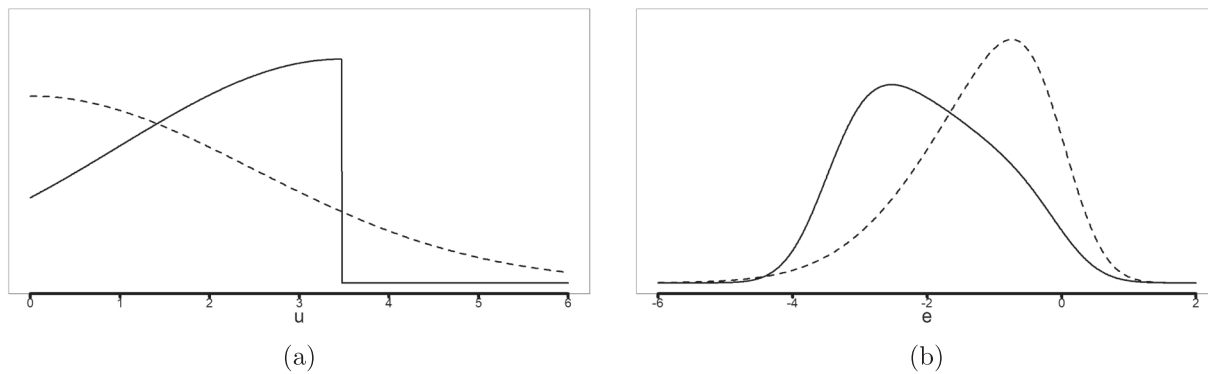


Fig. 2 Densities of u and e for $\gamma = 2.5$, i.e., correct skewness (dotted lines) and $\gamma = -2.5$, i.e., wrong skewness (solid lines); with $\sigma_v = 0.5$. **a** Density of u ; **b** Density of e

Hafner et al. (2018), we distinguish two cases for u_i which characterize the shape of distribution of composite errors $e_i = v_i - u_i$:

‘Correct’ skewness : (2)
 $u_i \sim N_{[0,\infty)}(0, \gamma^2), \quad \gamma > 0$

‘Wrong’ skewness : (3)
 $u_i \sim N_{[0,a_0|\gamma]}(a_0|\gamma, \gamma^2), \quad \gamma < 0$

The assumption in (2) is well-disseminated and describes “correct” skewness of u_i (and thus also e_i) because the density of u_i is strictly decreasing in $[0, \infty)$ (Kumbhakar and Lovell 2003).⁷ By contrast, “wrong” skewness is induced by (3), where $a_0 \approx 1.389$ is the non-trivial solution of $\frac{\phi(0)}{\Phi(0)} = a_0 + \frac{\phi(a_0) - \phi(0)}{\Phi(a_0) - \Phi(0)}$ and the density of u_i is strictly increasing and bounded in $[0, a_0|\gamma]$. It is worth highlighting that expectations of both u_i and e_i remain unaffected by the sign of skewness (Hafner et al. 2018). Thus, inefficiency variance and sign of skewness are directly related because $\gamma > 0$ ($\gamma < 0$) induces correct (“wrong”) skewness but $\mathbb{E}[u_i]$ and $\mathbb{E}[e_i]$ are not subject to the sign of γ . The density of $e_i = v_i - u_i$ is given by:

‘Correct’ skewness : (4)
 $g_e^+(e) = \frac{2}{\sigma} \phi\left(\frac{e}{\sigma}\right) \Phi\left(-\frac{e\gamma}{\sigma\sigma_v}\right),$

‘Wrong’ skewness : (5)
 $g_e^-(e) = \frac{1}{\sigma(\Phi(a_0) - \Phi(0))} \phi\left(\frac{e - a_0\gamma}{\sigma}\right) \left[\Phi\left(A_w + \frac{a_0\sigma}{\sigma_v}\right) - \Phi(A_w) \right],$
 $A_w = \frac{e - a_0\gamma}{\sigma} \frac{\gamma}{\sigma_v},$

⁷ In general, “correct” skewness in the production SF model means positive skewness of u_i and in consequence negative skewness of $e_i = v_i - u_i$ due to symmetry of v_i .

with $\sigma^2 = \gamma^2 + \sigma_v^2$ and $\int e g_e^+(e) de = \int e g_e^-(e) de$. Note that under “correct” skewness, $e \sim CSN(0, \sigma^2, -\frac{\gamma}{\sigma\sigma_v}, 0, 1)$, while as shown by Haschka and Wied (2022), under “wrong” skewness, it is $e \sim CSN_{1 \times 2}\left(a_0\gamma, \sigma^2, \begin{pmatrix} \gamma/\sigma \\ -\gamma/\sigma \end{pmatrix}, \begin{pmatrix} -a_0\sigma \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_v^2 & 0 \\ 0 & \sigma_v^2 \end{pmatrix}\right)$. The shape of inefficiency distribution under “correct” and “wrong” skewness is shown in Panel (a) of Fig. 2, and the corresponding distributions of composite errors in Panel (b). For $\gamma > 0$, the distribution of u (e) has positive (negative) skewness, whereas for $\gamma < 0$ its skewness is negative (positive). Note that composite error distributions in both cases are only determined by σ_v and γ . Accordingly, we can distinguish “correct” and “wrong” skewness without the necessity to identify further parameters (Hafner et al. 2018).⁸

3.2 Joint estimation using copulas

Let $F(z_1, \dots, z_K, e)$ and $f(z_1, \dots, z_K, e)$ be the joint distribution and the joint density of (z_1, \dots, z_K) and e , respectively. In practice, $F(\cdot)$ and $f(\cdot)$ are typically unknown and hence need to be estimated. Following Park and Gupta (2012), we adopt a copula approach to construct this joint density. The copula essentially captures dependence in the joint distribution of endogenous regressors and composed errors.

Let $\omega_{z,i} = (F_{z1}(z_{1i}), \dots, F_{zK}(z_{Ki}))'$ and $\omega_{e,i} = G(e_i; \sigma_v, \gamma)$ denote the margins $(\omega_{z,i}, \omega_{e,i})' \in [0, 1]^{K+1}$ based on a

⁸ The adopted one-sided distribution is parsimonious. However, other approaches that allow for a data-driven choice of correct or “wrong” skewness either involve the identification of multiple parameters that determine inefficiency distribution (see, e.g. Tsionas 2007, for Weibull inefficiency), or an a priori determination of the sign of skewness (COLS or MOLS). While Li (1996) argues that a one-sided error component with unbounded range always has a positive skewness, Johnson et al. (1995) shows that the two-parameter Weibull distribution can have positive and (small) negative skewness for specific parameter combinations.

probability integral transform. The F 's denote the respective marginal cumulative distributions functions of observed endogenous regressors and $G(e_i; \sigma_v, \gamma)$ is the cumulative distribution function of the CSN distribution for errors, which is subject to the sign of γ . Building on Tran and Tsionas (2015), we replace $F_1(z_{1i}), \dots, F_p(z_{pi})$ by their respective empirical counterparts in a first stage. Given observed samples of $z_{ji}, j = 1, \dots, p; i = 1, \dots, n$, we use the empirical cumulative distribution function of z_j , i.e., $\hat{F}_j = \frac{1}{n+1} \sum_{i=1}^n \mathbb{1}(z_{ji} \leq z_{0j})$.⁹

Using a Gaussian copula, $\hat{\xi}_{z,i} = (\Phi^{-1}(\hat{F}_{z1}(z_{1i})), \dots, \Phi^{-1}(\hat{F}_{zK}(z_{Ki})))'$, and $\hat{\xi}_{e,i} = \Phi^{-1}(\hat{G}(\hat{e}_i; \hat{\sigma}_v, \hat{\gamma}))$ follow a standard multivariate normal distribution of dimension $(K+1)$ with correlation matrix Ξ .¹⁰ Then, the joint density can be derived as

$$f(z_i, e_i) = \frac{1}{\sqrt{\det(\Xi)}} \exp\left(-\frac{1}{2} \begin{pmatrix} \hat{\xi}_{z,i} \\ \hat{\xi}_{e,i} \end{pmatrix}' (\Xi^{-1} - I) \begin{pmatrix} \hat{\xi}_{z,i} \\ \hat{\xi}_{e,i} \end{pmatrix}\right) \times g(e_i; \sigma_v, \gamma) \times \prod_{k=1}^K f_{zk}(\tilde{z}_{ki}), \quad (6)$$

where $\xi_{e,i}$ and $g(e_i; \sigma_v, \gamma)$ is again subject to “correct” or “wrong” skewness. The copula density in the first row links the error and all explanatory variables to encode information about the entire dependence in the model whereas densities in the second row describe marginal behavior. The marginal densities $f_{zk}(\tilde{z}_{ki})$ in (6) do not contain any parameter of interest and can be dropped when deriving the likelihood, since they enter as normalizing constants.

Before deriving the likelihood, we briefly discuss model identification. Under our setting, model identification requires

the distribution of endogenous regressors to be different from that of the composite error (for a more detailed discussion on identification issues, see Haschka 2022b; Park and Gupta 2012). Accordingly, the model is identified as long as γ is not zero (or very close to zero) and endogenous regressors are not normally distributed. However, model identification breaks down if both (i) $\gamma = 0$ (such that the composite error is normal) and (ii) endogenous regressors are normal. In this case, the joint distribution of endogenous regressors and composite error is multivariate normal, which implies that $\mathbb{E}[e|z]$ is a linear function, making it impossible to identify the linear effect δ without instrumental variable information (Haschka 2022b). Thus, external instrumental information is needed to provide model identification (Tran and Tsionas 2013). Consequently, the identification problem has important implications when $|\gamma| \rightarrow 0$. In this scenario, identification requires the endogenous regressors to be non-normally distributed. Therefore, in empirical applications, assessing the marginal distribution of endogenous regressors before estimation is a common approach in the empirical literature using copula-based identification (e.g., Datta et al. 2017; Haschka and Herwartz 2022; Papies et al. 2017).

To obtain a simultaneous choice of “correct” or “wrong” skewness that is determined by the sign of γ , we follow Haschka (2024) and use an indicator function for the likelihood. As an alternative to using indicator function in the likelihood, Hafner et al. (2018) argue that choice of “correct” or “wrong” skewness can be made a priori by inspecting skewness of the OLS residual. However, in our approach, the sign of γ is not predetermined but is instead estimated simultaneously with all other parameters. This approach is adopted because any prior determination of residual skewness could be influenced by (potential) endogeneity. Accordingly, we have

$$\begin{aligned} L(\theta|y, z, x) &\propto \mathbb{1}(\gamma > 0) \prod_{i=1}^n \frac{1}{\sqrt{\det(\Xi)}} \exp\left(-\frac{1}{2} \begin{pmatrix} \hat{\xi}_{z,i}^+ \\ \hat{\xi}_{e,i}^+ \end{pmatrix}' (\Xi^{-1} - I) \begin{pmatrix} \hat{\xi}_{z,i}^+ \\ \hat{\xi}_{e,i}^+ \end{pmatrix}\right) g^+(e_i; \sigma_v, \gamma) \\ &+ \mathbb{1}(\gamma < 0) \prod_{i=1}^n \frac{1}{\sqrt{\det(\Xi)}} \exp\left(-\frac{1}{2} \begin{pmatrix} \hat{\xi}_{z,i}^- \\ \hat{\xi}_{e,i}^- \end{pmatrix}' (\Xi^{-1} - I) \begin{pmatrix} \hat{\xi}_{z,i}^- \\ \hat{\xi}_{e,i}^- \end{pmatrix}\right) g^-(e_i; \sigma_v, \gamma) \\ &+ \mathbb{1}(\gamma = 0) \prod_{i=1}^n \frac{1}{\sqrt{\det(\Xi)}} \exp\left(-\frac{1}{2} \begin{pmatrix} \hat{\xi}_{z,i}^0 \\ \hat{\xi}_{e,i}^0 \end{pmatrix}' (\Xi^{-1} - I) \begin{pmatrix} \hat{\xi}_{z,i}^0 \\ \hat{\xi}_{e,i}^0 \end{pmatrix}\right) \phi(e_i; \sigma_v). \end{aligned} \quad (7)$$

⁹ The rescaling factor $1/(n+1)$ instead of $1/n$ ensures that the empirical cumulative distribution is well bounded in $(0, 1)$.

¹⁰ In general, any other copula that is capable of modeling multivariate dependency structures can also be used. According to Papadopoulos (2022), the Gaussian copula is most flexible and has many desirable properties. Furthermore, if the true dependence is different from what the Gaussian copula assumes, literature has demonstrated its robustness to capture various non-Gaussian dependencies (Becker et al. 2022; Haschka 2022b; Park and Gupta 2012); although the true dependence should not be nonparametric (Haschka 2022a) or asymmetric (Papadopoulos 2022).

To explicitly consider the case of only fully efficient firms, the likelihood also allows for $\gamma = 0$. Here, the marginal distribution of the errors is a normal distribution with mean zero and variance σ_v^2 , it is $\xi_{e,i}^0 = e_i/\sigma_v$. Note that our approach nests those by Hafner et al. (2018), Tran and Tsionas (2015), and Park and Gupta (2012). In case of exogeneity of all regressors, i.e. $\rho_k = 0 \forall k = 1, \dots, K$, the likelihood in (7) collapses to that in Hafner et al. (2018); in case of “correct” skewness, i.e.,

$\gamma > 0$, that is the traditional SF model, the likelihood collapses to that in Tran and Tsionas (2015); and in case of only fully efficient firms, i.e., $\gamma = 0$, it collapses to that in Park and Gupta (2012). Finally, the likelihood is logarithmised and maximized with respect to the vector of unknown parameters $\theta = (\beta, \delta, \sigma_v^2, \gamma, \text{vech}[\Xi])$, where $\text{vech}[\Xi] = (\rho_1, \dots, \rho_K)'$ stacks the lower diagonal elements of the correlation matrix Ξ into a column vector.¹¹

3.3 Bootstrap inference and testing under endogeneity

In their recent work, Breitung et al. (2023) argue that it remains unclear a priori whether the standard properties of ML estimation hold for joint estimation using copulas, and under which assumptions they may carry over. The problem of deriving precise statements about limiting properties in the presence of a nonparametrically generated regressor is highly non-trivial. This issue is common to all copula-based approaches that rely on joint estimation of errors and endogenous regressors with a priori estimated cumulative distribution functions (Haschka 2022b; Park and Gupta 2012; Tran and Tsionas 2015). Nonetheless, Papadopoulos (2022) and Tran and Tsionas (2015) conjecture that such estimators exhibit consistency and asymptotic normality in SFA settings (for further simulation-based evidence on asymptotic behavior, see Haschka 2022b; Haschka and Herwartz 2022). Although specific consistency (i.e., robustness) theory for the quasi-MLE for Gaussian copula-based models exists (Prokhorov and Schmidt 2009), it is not clear if these results are valid in the case of generated regressors. Accordingly, no general, i.e., theoretical, statement can be made about the consistency of the estimator.

Although ML inference is believed to be (asymptotically) valid if the identifying assumptions are satisfied (Breitung et al. 2023), we follow the literature on copula-based endogeneity-correction models (Haschka 2022b; Park and Gupta 2012) and recommend bootstrapping to obtain standard errors. However, simple bootstrapping can be problematic in stochastic frontier models for two reasons: (i) subsampling with replacement from the empirical CDF of the data can change the skewness of the distribution, and (ii) wild bootstrapping may fail to mimic the properties of the underlying conditional distribution of the error term. To address these issues, we adopt Algorithm #3 from Simar and Wilson (2009) and modify it for the copula

model. The difference between the algorithm we are using and that from Simar and Wilson (2009) arises in step 2, where we draw from the conditional distribution of composed errors given regressors using a copula representation instead of drawing from the marginal distribution.

Algorithm 1 Bootstrapping procedure for the SF copula model

1. Using pairs of the sample data $\{(y_i, \mathbf{x}_i, \mathbf{z}_i)\}_{i=1}^n$, maximize the log-likelihood in (7) to obtain estimates $\hat{\theta}$; recover $\hat{\sigma}_v, \hat{\gamma}$, and $\hat{\Xi}$.
2. For $i = 1, \dots, n$, draw pairs of e_i^* from $g_{e|\mathbf{z}, \mathbf{x}} = c(G_e(e; \hat{\sigma}_v, \hat{\gamma}), \hat{F}_1(x_1), \dots, \hat{F}_L(x_L), \hat{F}_1(z_1), \dots, \hat{F}_K(z_K); \hat{\Xi})$ $g_e^*(e; \hat{\sigma}_v, \hat{\gamma})$, where $G_e(g_e)$ is the cdf (pdf) of the CSN distribution given in (4) for the case of “correct”, and in (5) for “wrong” skewness.
3. Compute $y_i^* = \mathbf{x}_i' \hat{\beta} + \mathbf{z}_i' \hat{\delta} + e_i^*$.
4. Using the pseudo-data $\{(y_i^*, \mathbf{x}_i, \mathbf{z}_i)\}_{i=1}^n$, compute bootstrap estimates.
5. Repeat the steps 2.–4. B times to obtain estimates $\{(\hat{\theta}_b^*)\}_{b=1}^B$.

The bootstrap estimates $\{(\hat{\theta}_b^*)\}_{b=1}^B$ can be used to estimate confidence intervals for the parameters of the model (Simar and Wilson 2009). We follow the recommendation by Haschka and Herwartz (2022) and use a high number of bootstrap replications (say 1999) when trying to bootstrap standard errors and confidence intervals in copula-based endogeneity-correction frontier models.

Significance can then be assessed based on parameter estimates and (bootstrap) standard errors. This allows for an empirical detection of endogeneity and “wrong” skewness. Under exogeneity, the generated regressors do not enter the asymptotic distributions, and the test statistic for $H_0: \rho_k = 0$ follows an asymptotic standard normal distribution (Breitung et al. 2023). Furthermore, the usual approach based on a textbook t -statistic remains valid. Accordingly, these results imply that under exogeneity, the null hypothesis $H_0: \gamma = 0$ can be easily tested as the asymptotic distribution remains valid. By contrast, testing the null hypothesis $H_0: \gamma = 0$ under endogeneity is challenging because the asymptotic distribution is unknown. As shown by Breitung et al. (2023) for linear regression models with Gaussian outcomes (i.e., no SF specifications), the asymptotic distribution depends on unknown parameters (under endogeneity). We assume that in the case of an SF specification (non-zero γ), the asymptotic distribution is different from that derived in Breitung et al. (2023)—more precisely, it depends on other, unknown parameters. Accordingly, we refrain from proposing a test for this null hypothesis. Nevertheless, since (bootstrap) standard errors are valid (as will be shown in the simulations), we recommend determining if $\gamma = 0$ if an interval plus or minus two times its standard error does not cover zero.

¹¹ The likelihood function is continuous for fixed γ , but is not continuous in γ in the transition from a 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 usually do not recommend using gradient-based methods but rather the derivative-free simplex method for numerical optimization of Nelder and Mead (1965) that is applicable to non-differentiable functions.

With parameter estimates, technical inefficiency u_i can be predicted based on Jondrow et al. (1982):

‘Correct’ skewness :

$$\hat{u}_i = \hat{E}(u_i|e_i) = \frac{\sqrt{\hat{\gamma}^2 + \hat{\sigma}_v^2}(\hat{\gamma}/\hat{\sigma}_v)}{1 + (\hat{\gamma}/\hat{\sigma}_v)^2} \left[\frac{\phi((\hat{\gamma}/\hat{\sigma}_v)\hat{e}_i/\sqrt{\hat{\gamma}^2 + \hat{\sigma}_v^2})}{1 - \Phi((\hat{\gamma}/\hat{\sigma}_v)\hat{e}_i/\sqrt{\hat{\gamma}^2 + \hat{\sigma}_v^2})} - \frac{(\hat{\gamma}/\hat{\sigma}_v)\hat{e}_i}{\sqrt{\hat{\gamma}^2 + \hat{\sigma}_v^2}} \right] \quad (8)$$

‘Wrong’ skewness :

$$\hat{u}_i = \hat{E}(u_i|e_i) = \int_0^{a_0} |\hat{\gamma}| \exp\{-u_i\} f^-(u_i|e_i) du_i \quad (9)$$

Note that only in case of “correct” skewness predicted technical a has closed form expression, whereas the integral under “wrong” skewness has to be solved numerically (Hafner et al. 2018).

4 Monte Carlo simulations

To examine the finite sample properties of the proposed estimator, we conduct Monte Carlo experiments.¹² For comparison purposes, we also compute the MLE by Hafner et al. (2018), which considers both “correct” and “wrong” skewness but does not take endogeneity into account, an instrument-based GMM estimator (Tran and Tsionas 2013), and an IV-free copula estimator (Tran and Tsionas 2015), both of which account for endogenous regressors but assume correctly skewed inefficiency. In summary, our comparative investigation involves four estimators.

We consider two scenarios. First, we introduce endogeneity through correlation between explanatory variables and the two-sided noise component, following Tran and Tsionas (2015). In the second scenario, we consider the case where endogeneity results from correlation between the explanatory variable and the one-sided inefficiency component. In each scenario, we also generate an instrumental variable to implement the GMM estimator.

4.1 Correlation between regressors and idiosyncratic noise

Consider the following data generating process (DGP):

$$y_i = \beta x_i + \delta z_i + v_i - u_i \quad (10)$$

¹² In all simulations, the starting values for the proposed estimator are chosen as follows: for the idiosyncratic variance, the starting value is given by the empirical variance of the dependent variable, i.e., $\hat{\sigma}_v^2 = \text{Var}[y]$; for all other coefficients, we use 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.

$$z_i = \alpha s_i + \eta_i \quad (11)$$

$$u_i \sim \begin{cases} N_{[0,\infty)}(0, \gamma^2) & \text{if } \gamma > 0, \\ N_{[0,a_0|\gamma]}(a_0|\gamma, \gamma^2) & \text{if } \gamma < 0, \end{cases} \quad (12)$$

In (12), positive (negative) values of γ result in the distribution of u_{it} having positive (negative) skewness. Recall that both distributions have the same expectation. The random variables x_i and s_i are each generated independently as $\chi^2(2)$. To introduce endogeneity, the vector of errors $(v_i, \eta_i)'$ is generated by:

$$\begin{pmatrix} v_i \\ \eta_i \end{pmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}\right). \quad (13)$$

In our experiments, we fix $\alpha = \beta = \delta = 0.5$ and rescale $v_i = v_i/2$ such that the signal-to-noise ratio is $\gamma/\sigma_v = 2$. To assess scenarios under exogeneity and endogeneity, we set $\rho = \{0, 0.7\}$ and further distinguish $\gamma = \{-1, 1\}$ to introduce either negative or positive skewness. Finally, the sample size is $N = 750$ and we replicate each experiment 1000 times.

Estimation results for scenarios of exogeneity ($\rho = 0$) and endogeneity ($\rho = 0.7$) with either “correct” or “wrong” skewness are shown in Table 1. It is interesting to observe how endogeneity alters the skewness of the OLS residuals. Under exogeneity, assessing the sign of the skewness of the OLS residuals appear valid for detecting “wrong” skewness. However, the presence of endogenous regressors significantly complicates the detection of “wrong” skewness, as the OLS residuals indicate “correct” skewness. Specifically, when the true inefficiency term exhibits “wrong” skewness, the (average) skewness of the OLS residuals change from 0.0439 to -0.0276 when endogeneity is introduced. Therefore, under endogeneity, it is notable that the ML estimator fails to detect “wrong” skewness.

When the skewness of the inefficiency distribution is correctly specified (upper panel) and regressors are exogenous, all estimators are unbiased. In this scenario, ML is the method of choice as it yields the smallest standard deviations, while the remaining approaches (GMM, copula, proposed) are characterized by higher uncertainty as they unnecessarily account for endogeneity. However, when endogeneity is introduced, ML quickly deteriorates and becomes biased for all model parameters. While GMM, copula, and the proposed method perform equally well and remain unbiased, GMM estimates are most tightly centered around true values. This is because it is IV-based, and access to valid instruments strongly benefits efficiency (for similar findings, see Tran and Tsionas 2015).

Misspecifying the skewness of the inefficiency distribution (lower panel) has particularly detrimental effects on the GMM and copula estimators. In the absence of endogeneity ($\rho = 0$),

Table 1 Monte Carlo results for simulations using DGP in (10)–(13) under exogeneity ($\rho = 0$) and endogeneity ($\rho = 0.5$) with “correct” ($\gamma > 1$) and “wrong” skewness ($\gamma < 0$), respectively, and $N = 750$

			ML (Hafner et al. 2018)			GMM (Tran and Tsionas 2013)			Copula (Tran and Tsionas 2015)			Proposed				
Parameter			Mean	Sd	MSE	Mean	Sd	MSE	Mean	Sd	MSE	Mean	Sd	Se	MSE	
$\gamma = 1$	$\rho = 0$	β	0.5001	0.0211	0.0004	0.4999	0.0220	0.0005	0.4999	0.0204	0.0004	0.5018	0.0215	0.0199	0.0005	
		δ	0.5003	0.0304	0.0009	0.4998	0.0998	0.0099	0.4859	0.1442	0.0209	0.4952	0.1438	0.1449	0.0207	
		σ_v	0.4967	0.0620	0.0038	0.4977	0.0629	0.0038	0.4958	0.0615	0.0038	0.5011	0.0785	0.0801	0.0063	
		γ	0.9966	0.1044	0.0110	0.9967	0.1481	0.0220	0.9834	0.1822	0.0334	0.9836	0.1841	0.1833	0.0342	
		ρ	–	–	–	0.0099	0.1627	0.0264	0.0192	0.1822	0.0336	–0.0074	0.1821	0.1838	0.0332	
		$\mathbb{E}[u e]$	–	–	0.0672	–	–	0.0688	–	–	0.0691	–	–	–	–	0.0701
		Average skewness of the OLS residuals: –0.1350 (Q1: –0.2056; Q3: –0.0687)														
$\rho = 0.7$	β	0.4981	0.0179	0.0004	0.5058	0.0222	0.0005	0.5001	0.0221	0.0005	0.4971	0.0222	0.0223	0.0005		
	δ	0.7478	0.0296	0.0622	0.4950	0.0999	0.0100	0.5032	0.1457	0.0212	0.5037	0.1450	0.1431	0.0210		
	σ_v	0.2560	0.0501	0.0619	0.4980	0.0620	0.0038	0.4942	0.0632	0.0040	0.5016	0.0801	0.0788	0.0067		
	γ	1.311	0.0710	0.0980	0.9983	0.1488	0.0219	0.9909	0.1842	0.0344	0.9879	0.1861	0.1850	0.0348		
	ρ	–	–	–	0.4920	0.1661	0.0278	0.5029	0.1885	0.0356	0.4981	0.1871	0.1888	0.0350		
	$\mathbb{E}[u e]$	–	–	0.4299	–	–	0.0699	–	–	0.0695	–	–	–	–	0.0722	
	Average skewness of the OLS residuals: –0.2491 (Q1: –0.3138; Q3: –0.1787)															
$\gamma = -1$	$\rho = 0$	β	0.5000	0.0199	0.0004	0.4996	0.0208	0.0004	0.4979	0.0200	0.0004	0.5012	0.0201	0.0184	0.0004	
		δ	0.4985	0.0275	0.0007	0.3895	0.0903	0.0203	0.4011	0.1378	0.0287	0.4931	0.1759	0.1800	0.0310	
		σ_v	0.5018	0.0304	0.0013	0.5011	0.0298	0.0014	0.4891	0.0302	0.0016	0.5035	0.0533	0.0541	0.0041	
		γ	–0.9719	0.0790	0.0070	0.7794	0.1312	3.182	0.7395	0.1695	30.053	–0.9637	0.2351	0.2341	0.0566	
		ρ	–	–	–	0.1411	0.1645	0.0469	0.1240	0.1773	0.0469	–0.0055	0.2422	0.2433	0.0587	
		$\mathbb{E}[u e]$	–	–	0.0590	–	–	0.6922	–	–	0.7043	–	–	–	–	0.0761
		Average skewness of the OLS residuals: 0.0424 (Q1: 0.0067; Q3: 0.0780)														
$\rho = 0.7$	β	0.4995	0.0186	0.0003	0.4990	0.0178	0.0003	0.4993	0.0184	0.0003	0.5019	0.0193	0.0200	0.0004		
	δ	0.7469	0.0281	0.0619	0.1411	0.1422	0.7583	0.1400	0.1800	0.0619	0.5161	0.1918	0.1915	0.0371		
	σ_v	0.4140	0.0327	0.0085	0.4275	0.0302	0.0359	0.4140	0.0321	0.0085	0.5011	0.0337	0.0332	0.0013		
	γ	0.3615	0.0718	10.859	0.7318	0.1388	30.139	0.4166	0.2430	2.059	–0.9828	0.2299	0.2305	0.0532		
	ρ	–	–	–	0.7742	0.1590	0.2342	0.7747	0.1910	0.1122	0.5131	0.2314	0.2298	0.0537		
	$\mathbb{E}[u e]$	–	–	0.3676	–	–	0.7271	–	–	0.8292	–	–	–	–	0.0750	
	Average skewness of the OLS residuals: –0.0278 (Q1: –0.0599; Q3: 0.0006)															

The table shows the mean estimates, standard deviations of the respective coefficients, and mean squared errors. For each simulation experiment, the skewness of the OLS residuals was stored, and the table shows the average skewness as well as the first and third quantiles of the distribution in parentheses

both estimators yield biased slope coefficients. This finding is intriguing, as the misspecification of the error distribution due to “wrong” skewness is falsely identified as endogeneity, leading to biased estimates (for further simulation-based evidence, see Haschka and Herwartz 2022). Consequently, efficiency scores are also biased. In the case of endogeneity ($\rho = 0.7$), GMM and copula remain biased due to the misspecified error distribution, while ML is now also biased.

The proposed approach remains generally unaffected, being unbiased for all coefficients, and provides accurate assessments of efficiency. Therefore, the simultaneous consideration of the skewness sign, coupled with the use of a Gaussian copula to handle regressor endogeneity, proves adequate and allows for unbiased assessments.

4.2 Correlation between regressors and stochastic inefficiency

While most simulation studies introduce endogeneity as correlation between regressors and idiosyncratic noise (e.g. Tran and Tsionas 2013, 2015, refer to the setup in the previous subsection), we next assess whether the proposed approach can handle endogeneity stemming from the correlation between regressors and stochastic inefficiency. For this purpose, consider the following DGP:

$$y_i = \beta x_i + \delta z_i + v_i - u_i \tag{14}$$

$$v_i \sim N(0, 0.25) \tag{15}$$

Table 2 Monte Carlo results for simulations using DGP in eqs. (14)–(18)

Parameter	ML (Hafner et al. 2018)			GMM (Tran and Tsionas 2013)			Copula (Tran and Tsionas 2015)			Proposed			
	Mean	Sd	MSE	Mean	Sd	MSE	Mean	Sd	MSE	Mean	Sd	Se	MSE
$\gamma = 1$													
β	0.4956	0.0189	0.0004	0.5012	0.0225	0.0005	0.5021	0.0199	0.0004	0.5014	0.0229	0.0307	0.0005
δ	0.1815	0.0372	0.1028	0.8184	0.1415	0.1214	0.5240	0.1496	0.0230	0.5018	0.1477	0.1589	0.0218
σ_v	0.4511	0.0531	0.0051	0.4195	0.0718	0.0116	0.4911	0.0530	0.0029	0.5028	0.0790	0.0799	0.0070
γ	0.9657	0.0858	0.0085	0.8195	0.1501	0.0551	1.008	0.0820	0.0067	1.200	0.1871	0.1802	0.0750
ρ	–	–	–	–0.6894	0.1675	10.690	–0.3083	0.1417	0.0201	–0.3099	0.1870	0.1864	0.0351
$\mathbb{E}[u e]$	–	–	0.3734	–	–	0.4208	–	–	0.0703	–	–	–	0.0710
Average skewness of the OLS residuals: –0.1043 (Q1: –0.1621; Q3: –0.0439)													
$\gamma = -1$													
β	0.4911	0.0187	0.0004	0.4995	0.0224	0.0005	0.4976	0.0184	0.0184	0.5017	0.0312	0.0301	0.0010
δ	0.2922	0.0345	0.0444	0.3891	0.1433	0.0328	0.6454	0.1471	0.0428	0.4891	0.1500	0.1489	0.0226
σ_v	0.3926	0.0504	0.0141	0.0984	0.0511	0.0123	0.8365	0.0412	0.1349	0.5028	0.0442	0.0458	0.0027
γ	0.9438	0.0814	0.0098	0.7844	0.1401	0.0661	0.9889	0.0780	0.0062	–1.027	0.2070	0.2081	0.0435
ρ	–	–	–	–0.1844	0.1500	0.0232	–0.3138	0.0591	0.0142	–0.2164	0.2241	0.2250	0.0503
$\mathbb{E}[u e]$	–	–	0.2239	–	–	0.4223	–	–	0.2561	–	–	–	0.0777
Average skewness of the OLS residuals: –0.0657 (Q1: –0.1237; Q3: 0.0049)													

For further description, see Table 1

Again, the random variable x_i is generated independently as $\chi^2(2)$. To introduce correlation between z_i and u_i , we draw:

$$\begin{pmatrix} \eta_i \\ \xi_{ui} \end{pmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & 0.7 \\ 0.7 & 1 \end{bmatrix}\right). \quad (16)$$

$$z_i = \eta_i + s_i \quad (17)$$

$$u_i = \begin{cases} F_{N_{[0,\infty)}(0,\gamma^2)}^{-1}(\Phi(\xi_{ui})) & \text{if } \gamma > 0, \\ F_{N_{[0,a_0|\gamma|]}(a_0|\gamma|,\gamma^2)}^{-1}(\Phi(\xi_{ui})) & \text{if } \gamma < 0, \end{cases} \quad (18)$$

$F_{N_{[0,\infty)}(0,\gamma^2)}^{-1}$ and $F_{N_{[0,a_0|\gamma|]}(a_0|\gamma|,\gamma^2)}^{-1}$ are inverse cumulative distribution function of the half normal and truncated normal distributions, respectively. Finally, we keep all other parameters (including the signal-to-noise ratio) and variables unchanged when comparing with the simulations in the previous section.

Simulation results are presented in Table 2. Similar to the findings in the previous section, the OLS residuals consistently exhibit negative skewness in the presence of endogeneity. This observation is intriguing because, according to the assumed data generating process (DGP) with $\gamma = -1$ (lower panel), a positive skewness is expected. Hence, it is unsurprising that the ML estimator fails to detect the “wrong” skewness.

While stylized ML estimation encounters significant bias stemming from endogeneity, it is worth noting that the GMM estimator also struggles to address this bias effectively. This is due to the assumption underlying the GMM estimator,

which posits that endogeneity arises from correlation with idiosyncratic noise rather than inefficiency (Tran and Tsionas 2013). An intriguing observation is that although the GMM estimator correctly identifies endogeneity, it attributes it incorrectly to correlation with idiosyncratic noise. As a result, the estimates obtained are biased in the opposite direction compared to those of the ML estimator.

Under correct skewness, the copula estimator by Tran and Tsionas (2015) yields unbiased estimates. This indicates that the Gaussian copula-based endogeneity correction, modeling the joint distribution of endogenous regressors and composite errors, effectively handles endogeneity arising from correlation with idiosyncratic noise or inefficiency. Our simulations thus validate the claims made by Tran and Tsionas (2015), as this scenario had not been previously explored in research. However, under the “wrong” skewness assumption, this copula estimator exhibits severe bias and fails to identify the true coefficients. In contrast, the proposed estimator provides unbiased estimates regardless of the skewness sign, with standard errors showing only marginal changes compared to simulations where endogeneity resulted from correlation with idiosyncratic noise (see Table 1).

5 Estimating firm efficiency in Vietnam

We provide empirical evidence of the applicability of the proposed approach by evaluating firm efficiency using data on Vietnamese firms in 2015 obtained from the Vietnam

Table 3 Summary statistics

Variable	N. obs.	Mean	Median	Std. Dev.	IQR	Min	Max
revenues	16,474	21,981	2658	82,149	9050	9	1,226,094
wages	16,474	2387	466.3	9588	941.6	3	359,494
assets	16,474	9858	1594	44,189	3457	2.5	1,373,647

All variables are measured in million Vietnamese Dong. Revenues denotes total revenue from sales and supplying of services, wages denotes total amount spent on wages, and assets denote total value of fixed assets

Enterprise Survey (VES). The dataset consists of 16,641 observations. Conducted annually by the General Statistics Office (GSO) of Vietnam, the VES is a nationally representative survey that includes all firms with 30 or more employees, as well as a representative sample of smaller firms (O’Toole and Newman 2017).¹³ Focusing on firms within a single (developing) country allows us to best unravel endogenous interrelations and their implications for understanding “wrong” skewness empirically. This approach helps us avoid difficulties associated with time dependencies or panel analysis, as additional endogeneities channeled through unobserved heterogeneity are ruled out. To assess firm performance, we adopt common specifications employed in related literature by relating firm revenues to wages and assets (Haschka et al. 2023, 2021). Following Haschka et al. (2021) and related studies using VES data, we first remove outlier observations from the sample. Observations are excluded if revenues fall outside the 99.5th percentile or the 0.5th percentile of their distributions. This process leaves us with a sample size of 16,474. Descriptive statistics of the continuous variables involved are presented in Table 3.

By considering each firm as a single producer, we use a log-linear Cobb-Douglas production function and specify the following model:

$$\log \text{revenue}_i = \alpha + \delta_1 \log \text{wages}_i + \delta_2 \log \text{assets}_i + \mathbf{x}'_i \boldsymbol{\beta} + v_i - u_i, \quad (19)$$

where $i = 1, \dots, 16,474$ denotes firms, and $v_i \sim N(0, \sigma_v^2)$ is idiosyncratic noise. The firms may be very different according to the sector in which they operate, and unobserved regional effects such as local subsidies or advantages of location may also have a meaningful effect on output generation. Accordingly, the vector \mathbf{x}_i contains sector-specific and regional dummy variables to tease out heterogeneity in the error term and to immunize our analysis to adverse effects stemming from unobserved cross-industry or regional heterogeneities as another potential channel of endogeneity.

¹³ Each firm in the VES dataset is classified according to the Vietnam Standard Industrial Code (VSIC). For more information, see <https://vietnamcredit.com.vn/products/industries>.

Our specification diverges from related models in the following two directions. First, we allow stochastic inefficiency to vary over i and consider both “correct” and “wrong” skewness by means of a data-driven choice of distribution of u_i , i.e.,

$$u_i = 0 \quad \text{or} \quad u_i \sim N_{[0, \infty)}(0, \gamma^2) \quad \text{or} \quad u_i \sim N_{[0, a_0|\gamma]}(a_0|\gamma, \gamma^2). \quad (20)$$

The latter case has yet not been considered in empirical development literature and thus offers a novel perspective to unravel structural inefficiencies in firm performance in Vietnam. Note that although we label it as “data-driven choice”, the sign of γ is not estimated a priori but rather simultaneously with all other parameters because any a priori determination of residual skewness would be subject to (potential) endogeneity. In this regard, we consider the possibility of correlated production inputs with composed errors, denoted as $e_i = v_i - u_i$. Endogeneity of inputs can arise due to their correlation with v_i , with u_i , or both. The presence of omitted variables in the production function, such as subsidies or governmental grants that are large enough to have meaningful effects may lead to correlation with v_i . Furthermore, if producers possess prior knowledge of potential inefficiencies in output generation, they are likely to adjust their inputs accordingly (Haschka and Herwartz 2022). As these adjustments are unobserved by the analyst, they introduce correlation with u_i .

5.1 Estimator setup

Before estimation, we assess whether the identifying assumptions for the proposed estimator are met. Following Papadopoulos (2022), we first test for the multivariate normality of the endogenous regressors under a copula-based transformation to verify the adequacy of the Gaussian copula. The Doornik-Hansen test yields a p value of $p = 0.3014$, indicating that the Gaussian copula assumption is not violated. Next, we examine whether the marginal distribution of the endogenous regressors deviates sufficiently from normality. Figure 3 displays the distribution of the log-transformed endogenous variables. It is evident that the distributions are skewed and exhibit excess kurtosis. Therefore, this deviation should provide ample information

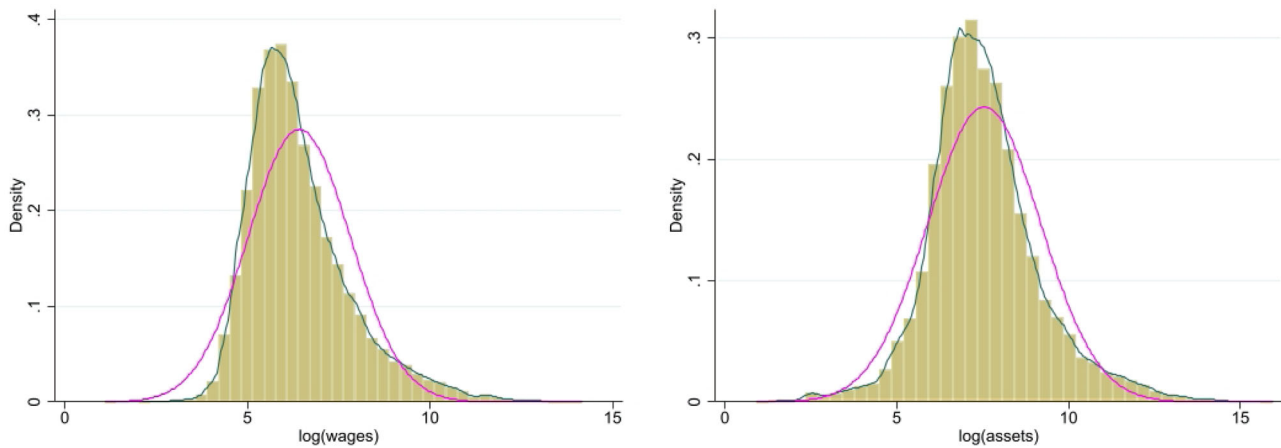


Fig. 3 Histograms and density plots of the explanatory variables $\log(\text{wages})$ and $\log(\text{assets})$. The magenta curves show fitted normal distributions. In addition, Cramér-von Mises tests were performed to test for normality, which yielded p values of $p < 0.001$ in both cases

for model identification even in the absence of skewness in the error distribution.

The model encompasses numerous unknown parameters (because of sector and regional dummies), and the likelihood function exhibits discontinuity in γ around values of zero skewness. Since numerical optimization techniques are employed (Nelder and Mead 1965), it is imperative to ensure that the optimizer avoids getting trapped in local maxima, particularly given the absence of theoretical insights into the continuity of likelihood functions in copula-based endogeneity correction models. Consequently, there is no guarantee that the optimizer consistently converges to the global maximum, and its performance may be influenced by the selection of starting values.

To address these challenges, we employ a grid-based approach to search for optimal starting values for numerical optimization. The process unfolds as follows: we estimate the model in (19) using stylized least squares estimation to derive estimates for the regression coefficients and residual variance. The grid is constructed around these estimates, spanning \pm ten times their corresponding standard errors, and consists of 100 values for each parameter. For the skewness parameter, the grid extends from -5 to 5 in increments of 0.1 . We initialize the optimizer from each of these potential starting points and select the model with the maximum (log) likelihood value. Finally, bootstrapping is employed on the optimal model to derive standard errors.

We conduct a comparative analysis between the proposed estimator, which accommodates endogeneity of inputs and both “correct” and “wrong” skewness of inefficiency, and the maximum likelihood estimator (MLE) introduced by Hafner et al. (2018). The MLE also considers “correct” and “wrong” skewness but does not address endogeneity. Furthermore, we compare it to the instrument-based generalized method of moments (GMM) estimator

(Tran and Tsionas 2013) and the IV-free copula estimator (Tran and Tsionas 2015), both of which handle endogenous regressors but assume correctly skewed inefficiency. For IV-based GMM estimation, we utilized one-year lagged assets and one-year lagged wages as instruments, following a methodology similar to previous studies (Haschka and Herwartz 2022). However, it is important to note that such internal instrumentation may suffer from weak instruments and might not be entirely suitable for addressing endogeneity.

5.2 Returns to scale and Vietnamese growth strategies

The estimation results are presented in Table 4. The proposed estimator reveals a negatively skewed inefficiency, whereas the remaining estimators all indicate “correct” skewness. Furthermore, all employed estimators consistently highlight human capital as the primary driver of firm performance in Vietnam. This is evident from the substantially higher coefficient attached to (log) wages compared to the coefficient attached to (log) assets, representing gross fixed capital formation. Accounting for endogeneity through GMM, copula, and the proposed estimator reduces this difference. It is worth noting that both GMM and copula estimators may still be affected by remaining endogeneity when “wrong” skewness is present, while GMM may face additional challenges due to weak instrumentation.

In examining the outcomes of the different estimators on the production function, distinct patterns in returns to scale (RTS) emerge. The ML estimator reveals increasing returns to scale, as the sum of coefficients is significantly greater than 1 ($RTS_{ML} = 1.153$, $CI = (1.11, 1.19)$). Conversely, both the GMM and copula estimators indicate significantly

Table 4 Estimation results using MLE (Hafner et al. 2018), GMM (Tran and Tsionas 2013), copula (Tran and Tsionas 2015), and the proposed estimator

	MLE		GMM		Copula		Proposed	
	Est.	SE	Est.	SE	Est.	SE	Est.	SE
<i>Const</i>	1.140	0.0625	0.9814	0.0752	1.055	0.0800	1.241	0.0812
<i>log wages</i>	0.9102	0.0114	0.6211	0.0198	0.5994	0.0201	0.6493	0.0210
<i>log assets</i>	0.2428	0.0086	0.2738	0.0188	0.2916	0.0199	0.3413	0.0213
σ_v	0.5918	0.0142	0.6561	0.0295	0.6290	0.0308	0.8993	0.0301
γ	0.6345	0.0310	0.8181	0.0393	0.8544	0.0409	-0.8803	0.0404
$\rho_{e,\log wages}$					0.3310	0.0391	0.2510	0.0404
$\rho_{e,\log assets}$					0.1822	0.0365	0.2442	0.0361
Mean Efficiency	0.7985		0.4252		0.4516		0.6552	
Sector Dummies	Yes		Yes		Yes		Yes	
Regional Dummies	Yes		Yes		Yes		Yes	

Standard errors of the copula-based estimators (copula and proposed) are obtained by means of bootstrap procedures with 1999 replications. Efficiency scores are calculated using the estimator by Jondrow et al. (1982). The skewness of the OLS residuals is -0.0981

decreasing returns to scale ($RTS_{GMM} = 0.8949$, $CI = (0.818, 0.972)$; and $RTS_{copula} = 0.891$, $CI = (0.8110, 0.971)$). This suggests that, on average, as firms increase their inputs, the growth rate of output diminishes. In contrast, the proposed estimator paints a different picture, depicting a scenario of constant returns to scale. This is visible as the sum of coefficients in the production function is roughly 1 ($RTS_{Proposed} = 0.9906$, $CI = (0.906, 1.08)$).¹⁴ While increasing RTS suggest that it should be easy for firms to scale up, decreasing RTS urges firms to assess their expansion strategies critically, as indiscriminate scaling might not yield proportional increases in output, and considerations for optimizing resource allocation and operational efficiency are paramount. Constant returns to scale are indicative of a more consistent and predictable production process, allowing firms to plan and allocate resources with greater confidence.

The question that now arises is which results seem most economically feasible for the case of Vietnam. Vietnam stood out as the sole emerging economy in Southeast Asia to avoid recession in 2009 amidst the global crisis. Moreover, Vietnam has demonstrated sustained growth rates over the past few decades (Cling et al. 2010). However, explaining increasing returns to scale would be difficult given the predominance of small firms in the dataset (O'Toole and Newman 2017). Although there is significant growth potential in the Vietnamese economy (Bai et al. 2019), the government prioritizes achieving stable and consistent economic growth over pursuing rapid growth at the expense of stability (Nguyen et al. 2018). This

¹⁴ Regarding the estimated RTS obtained by the proposed estimator, one can see that if a firm simultaneously increases its wages and assets by 1%, it can expect .9906% more revenues. Since this value is not significantly different from 1, we can conclude that returns to scale are constant.

perspective favors constant returns to scale, as it means that the government's economic goals can be attained while maintaining stability (Nghiem Tan et al. 2021). This suggests that MLE, GMM, and copula estimators may be flawed due to (remaining) endogeneity.

5.3 Endogeneity bias and efficiency levels

Additional evidence in favor of endogeneity is provided by significant estimates of correlations between production inputs and errors when using copula and the proposed estimators. Specifically, the correlation coefficients are estimated as $\hat{\rho}_{e,\log wages} = 0.2510$ and $\hat{\rho}_{e,\log assets} = 0.2442$ for the proposed estimator. These metrics directly reflect the interdependence and provide valuable economic information by allowing an evaluation of how firms adapt to fluctuations in inefficiency or random disturbances, on average (Haschka and Herwartz 2022). Substantial positive correlation estimates suggest notable adjustments in inputs in response to implicit shifts in production technology or idiosyncratic shocks. It seems intuitive for both correlation estimates to be positive. For example, adverse external technological shocks are likely to diminish efficiency, indirectly leading to reduced output. At the same time, an increase in production inputs becomes necessary to maintain the output level.

The distribution of efficiency scores is shown in Fig. 4. Considering firm efficiency, we find rather high mean firm efficiency when using MLE, with an average score of 0.8785. These results initially seem plausible, because it is in line with other efficiency levels documented in the literature (Le and Harvie 2010; Le et al. 2018; Nguyen et al. 2018; Tran et al. 2008; Vu 2003). However, it should be mentioned that none of these studies consider potential regressor-endogeneity. Accounting for endogeneity while

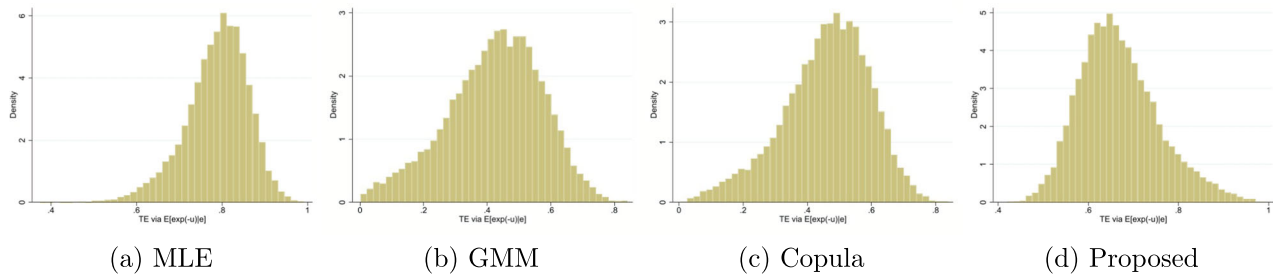


Fig. 4 Distribution of efficiency scores obtained by the four estimators via $TE_i = \mathbb{E}[-\exp\{\hat{u}_i\}]$ (see Eqns. (8) and (9)). **a** MLE; **b** GMM; **c** Copula; **d** Proposed

assuming “correct” skewness through GMM and copula estimators substantially decreases the mean efficiency scores to 0.4252 (GMM) and 0.4516 (copula), respectively. While these values seem very low, the proposed approach reveals a mean efficiency of 0.6552, still indicating a considerable shortfall of about 0.35% from maximum feasible output of Vietnamese firms. This is in line with Haschka et al. (2023), who also consider potential endogeneity and find similarly low efficiency levels. Important factors identified by the literature on low efficiency scores in Vietnam are corruption and the level of local financial development (Haschka et al. 2023, 2021). Previous studies have highlighted a direct correlation between corruption and inefficiency (Nguyen and Van Dijk 2012; Rand and Tarp 2012), as well as a negative impact of higher local financial development on firm efficiency in Vietnam (O’Toole and Newman 2017). These findings align with broader research indicating that while financial development tends to bolster technical efficiency in highly efficient economies, its effects are diminished or even adverse in less-efficient ones (Arestis et al. 2006; Rioja and Valev 2004).

Another possible explanation for low-efficiency levels is the connection with the adjustment of inputs, i.e., as a reflection of endogeneity (Haschka and Herwartz 2022). If firms are aware of their own low inefficiency levels, and increase their inputs according to that, a positive (unobserved) input-inefficiency dependence might be present. Since the simulations show that a positive correlation leads to an overestimation of the efficiency levels (see also Tran and Tsionas 2015), and the proposed estimator indicates a positive dependence between inputs and errors, this could provide another explanation for the actually lower efficiency levels. A positive sign of both correlation estimates is intuitively reasonable. For instance, (adverse) external technological shocks are likely to reduce efficiency and result indirectly in less output. At the same time, the production inputs have to be simultaneously increased to retain the output level. Exemplifying such shocks, one might notice environmental regulations of the Vietnamese government (Ho 2015). Consequently, stronger production

restrictions or sharper regulations might be considered as potential manifestations of endogenous technological shocks.

While the OLS residuals point to correct skewness and MLE also favors the traditional SF specification, the proposed estimator indicates the presence of “wrong” skewness after accounting for endogeneity. Although MLE can also detect “wrong” skewness, the presence of endogenous regressors likely hinders that. The pronounced difference in the efficiency scores obtained by GMM, copula, and the proposed estimator is interesting insofar as the latter indicates “wrong” skewness, but at the same time delivers higher efficiencies. We assume that since GMM and copula cannot deal with skewness issues, this was falsely identified as endogeneity, and the efficiency values are therefore underestimated (as could also be shown in the simulations).

5.4 Competition and market characteristics

The insights offered by the proposed estimator provide evidence supporting the presence of endogenous regressors and “wrong” skewness. While the former has already been emphasized in the existing empirical development literature, the latter has not yet been acknowledged. While correct skewness coupled with low efficiency scores would probably be accepted, “wrong” skewness gives reason to question the market conditions. As mentioned in Section 2, we argue that “wrong” skewness contradicts the assumption of a competitive market situation. Yet, unless alternative explanations for the “wrong” skewness can be ruled out, such as a poor sample or asymmetry in idiosyncratic noise, relying solely on an economic rationale may lack persuasiveness. The notion that “wrong” skewness is attributed to bad luck with the sample becomes less plausible given the substantial size of our dataset, which is representative for Vietnam (O’Toole and Newman 2017). While it would be conceivable that asymmetry exists in the two-sided noise term or this term may be correlated with inefficiency, a shift in the model specification towards asymmetric two-sided distributions or a dependency between inefficiency and noise would necessitate a sound economic rationale.

High efficiencies combined with the “wrong” skewness revealed by the proposed estimator, collectively cast doubt on the prevailing market conditions and raise concerns about competition levels. These findings might be attributed to a lack of incentives for firms to optimize their efficiency, resulting in many firms lacking the pressure to improve their operations. This situation can be attributed to various factors, despite the already mentioned corruption in Vietnam, burdens imposed by the communist regime, which hinder the emergence of liberal and competitive market conditions (Rand and Tarp 2012; Sahut and Teulon 2022), or a lack of market orientation of many companies (Evangelista et al. 2013). Despite extensive reforms in this regard, their effectiveness appears to be limited (Gupta et al. 2014; Tran et al. 2008). On the one hand, Vietnam established a central committee dedicated to combating corruption and enacted an anti-corruption law in 2005, followed by ratification of the UN Convention Against Corruption in 2009. Despite these proactive measures, the country’s standing in the Corruption Perception Index of 2018 remained relatively low, with a ranking of 117th among 180 countries. This position represented a decline of ten places compared to the previous year,¹⁵ indicating ongoing challenges in addressing corruption effectively. On the other hand, a lack of market orientation of firms in Vietnam is often attributed to the influence of cultural, economic, and institutional characteristics. As noted by Evangelista et al. (2013), market orientation can have a significant impact on business efficiency, when its adoption is influenced by a combination of internal organizational dynamics and external market forces. Therefore, further policy interventions are necessary to provide firms with incentives to optimize their processes and enhance efficiency since the market alone fails to generate adequate incentives in this regard. Specifically, steps to combat corruption are already being taken in the right direction, while incentives to strengthen the market orientation of Vietnamese producers could prove advantageous and contribute further to the overall improvement of business processes.

6 Conclusion

Under the traditional production SF specification, composite errors are assumed to have negative skewness. Violations of this assumption, commonly termed “wrong” skewness, have been highlighted in the literature on SF analysis (Choi et al. 2021; Curtiss et al. 2021; Daniel et al. 2019). While earlier discussions attributed such skewness issues to dataset peculiarities like small or poor samples (Almanidis and Sickles 2011; Hafner et al. 2018; Simar and

Wilson 2009), recent studies increasingly delve into economic rationales (see Papadopoulos and Parmeter 2023, for a review). We contribute to this discourse by outlining that when assuming that “wrong” skewness stems from the inefficiency term, it might indicate a lack of market incentives, leading producers to perceive no need to address their inefficiencies.

While various methodological approaches exist to address skewness issues empirically (e.g., Hafner et al. 2018), our work contributes to the economic explanation of “wrong” skewness. However, existing studies have not considered the potential endogeneity of regressors. Endogeneity of production inputs is likely if firms adjust their resource allocation based on their own inefficiency levels, which may go unnoticed by the analyst. In the presence of endogeneity, examining firm processes becomes challenging for analysts, making it even more difficult to uncover underlying market characteristics.

Against this background, the methodological scope of this paper is to propose an approach for estimating SF models while simultaneously identifying inefficiency skewness, considering potential endogeneity of regressors. Adapting the approach by Park and Gupta (2012), we utilize a Gaussian copula function to construct the joint distribution of endogenous regressors and composite errors, capturing mutual dependency without relying on instrumental variables. Our model distinguishes between “correct” and “wrong” skewness in one model without imposing a priori sign restrictions on inefficiency skewness or needing to identify additional parameters determining skewness. We assess the finite sample behavior of the proposed approach through Monte Carlo simulations.

We analyze the determinants of firm performance in Vietnam using data from 16,474 unaffiliated firms in 2015. Our findings reveal three key insights. Firstly, endogeneity significantly affects firm productivity estimation, indicating constant returns to scale rather than increasing returns. This finding aligns with the government’s goal for steady growth. Secondly, endogeneity complicates the detection of “wrong” skewness, leading to lower efficiency estimates. Lastly, evidence of low efficiency and “wrong” skewness suggests a lack of competitiveness due to factors like corruption and government constraints. Policy interventions are crucial to incentivize firms to optimize processes and improve efficiency.

Data availability

The dataset generated during the current study is not publicly available as it contains proprietary information that the authors acquired through a license. Information on how to obtain it and reproduce the analysis is available from the corresponding author on request.

¹⁵ <https://www.transparency.org/cpi2018>.

Funding Open Access funding enabled and organized by Projekt DEAL.

Compliance with ethical standards

Conflict of interest The author declares no competing interests.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Aigner D, Lovell CK, Schmidt P (1977) Formulation and estimation of stochastic frontier production function models. *J Econom* 6:21–37
- Almanidis P, Sickles R (2011) The skewness issue in stochastic frontiers models: fact or fiction? In van Keilegom I, Wilson PW (eds) *Exploring research frontiers in contemporary statistics and econometrics*, Springer, Berlin/Heidelberg, DE, p 201–227
- Amsler C, Prokhorov A, Schmidt P (2014) Using copulas to model time dependence in stochastic frontier models. *Econom Rev* 33:497–522
- Amsler C, Prokhorov A, Schmidt P (2016) Endogeneity in stochastic frontier models. *J Econom* 190:280–288
- Amsler C, Prokhorov A, Schmidt P (2017) Endogenous environmental variables in stochastic frontier models. *J Econom*. 199:131–140
- Amsler C, Prokhorov A, Schmidt P (2021) A new family of copulas, with application to estimation of a production frontier system. *J Prod Anal* 55:1–14
- Amsler C, Schmidt P (2021) A survey of the use of copulas in stochastic frontier models. In Parmeter C, Sickles RC (eds) *Advances in efficiency and productivity analysis*, Springer, p 125–138
- Arestis P, Chortareas G, Desli E (2006) Financial development and productive efficiency in OECD countries: an exploratory analysis. *Manchester Sch* 74:417–440
- Badunenko O, Henderson DJ (2024) Production analysis with asymmetric noise. *J Prod Anal*. 61:1–18
- Bai J, Jayachandran S, Malesky EJ, Olken BA (2019) Firm growth and corruption: empirical evidence from Vietnam. *Econ J* 129:651–677
- Battese GE, Coelli TJ (1995) A model for technical inefficiency effects in a stochastic frontier production function for panel data. *Empir Econ* 20:325–332
- Becker J-M, Proksch D, Ringle CM (2022) Revisiting Gaussian copulas to handle endogenous regressors. *J Acad Mark Sci* 50:46–66
- Bonanno G, De Giovanni D, Domma F (2017) The “wrong skewness” problem: a re-specification of stochastic frontiers. *J Prod Anal* 47:49–64
- Bonanno G, Domma F (2022) Analytical derivations of new specifications for stochastic frontiers with applications. *Mathematics* 10:3876
- Breitung J, Mayer A, Wied D (2023) Asymptotic properties of endogeneity corrections using nonlinear transformations (March 3, 2023). Available at arXiv. <https://arxiv.org/abs/2207.09246>
- Carree MA (2002) Technological inefficiency and the skewness of the error component in stochastic frontier analysis. *Econ Lett* 77:101–107
- Carta A, Steel MF (2012) Modelling multi-output stochastic frontiers using copulas. *Comput Stat Data Anal* 56:3757–3773
- Centorrino S, Pérez-Urdiales M (2023) Maximum likelihood estimation of stochastic frontier models with endogeneity. *J Econom* 1:82–105
- Choi K, Kang HJ, Kim C (2021) Evaluating the efficiency of Korean festival tourism and its determinants on efficiency change: parametric and non-parametric approaches. *Tour Manag* 86:104348
- Cincera M (1997) Patents, R&D, and technological spillovers at the firm level: Some evidence from econometric count models for panel data. *J Appl Econom* 12:265–280
- Cling J-P, Chi NH, Razafindrakoto M, Roubaud F (2010) How deep was the impact of the economic crisis in Vietnam? A focus on the informal sector in Hanoi and Ho Chi Minh City. Washington, DC, World Bank
- Cohen B, Winn MI (2007) Market imperfections, opportunity and sustainable entrepreneurship. *J Bus Ventur* 22:29–49
- Curtiss J, Jelínek L, Medonos T, Hruška M, Hüttl S (2021) Investors' impact on Czech farmland prices: a microstructural analysis. *Eur Rev Agric Econ* 48:97–157
- Daniel BC, Hafner CM, Simar L, Manner H (2019) Asymmetries in business cycles and the role of oil prices. *Macrocon Dyn* 23:1622–1648
- Das A (2015) Copula-based stochastic frontier model with auto-correlated inefficiency. *Centr Eur J Econom Model Econom* 7:111–126
- Datta H, Ailawadi KL, Van Heerde HJ (2017) How well does consumer-based brand equity align with sales-based brand equity and marketing-mix response? *J Mark* 81:1–20
- Dominguez-Molina JA, González-Farías G, Ramos-Quiroga R (2003) Skew-normality in stochastic frontier analysis. *Comun Téc. No. 1* 3–18
- Ehrenfried F, Holzner C (2019) Dynamics and endogeneity of firms' recruitment behaviour. *Labour Econ* 57:63–84
- El Mehdi R, Hafner CM (2014) Inference in stochastic frontier analysis with dependent error terms. *Math Comput Simul* 102:104–116
- Evangelista F, Thuy PN et al. (2013) Does it pay for firms in Asia's emerging markets to be market-oriented? Evidence from Vietnam. *J Bus Res* 66:2412–2417
- González-Farías G, Dominguez-Molina JA, Gupta A (2004) The closed skew normal distribution. In Genton, M (ed) *Skew elliptical distributions and their applications: a journey beyond normality*, Chapman and Hall/CRC, p 25–42
- Green A, Mayes D (1991) Technical inefficiency in manufacturing industries. *Econ J* 101:523–538
- Griffiths WE, Hajargasht G (2016) Some models for stochastic frontiers with endogeneity. *J Econom* 190:341–348
- Gupta R, Yang J, Basu PK (2014) Market efficiency in emerging economies—case of Vietnam. *Int J Bus Glob* 13:25–40
- Haschka RE (2024) Endogeneity in stochastic frontier models with “wrong” skewness: copula approach without external instruments. *Statistical Methods & Applications* (forthcoming). <https://doi.org/10.1007/s10260-024-00750-4>
- Hafner CM, Manner H, Simar L (2018) The “wrong skewness” problem in stochastic frontier models: a new approach. *Econom Rev* 37:380–400

- Haschka RE (2022) Bayesian inference for joint estimation models using copulas to handle endogenous regressors. Available at SSRN. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4235194
- Haschka RE (2022) Handling endogenous regressors using copulas: a generalisation to linear panel models with fixed effects and correlated regressors. *J Mark Res* 59:860–881
- Haschka RE, Herwartz H (2020) Innovation efficiency in European high-tech industries: evidence from a Bayesian stochastic frontier approach. *Res Policy* 49:104054
- Haschka RE, Herwartz H (2022) Endogeneity in pharmaceutical knowledge generation: an instrument-free copula approach for Poisson frontier models. *J Econ. Manag Strat* 31:942–960
- Haschka RE, Herwartz H, Silva Coelho C, Walle YM (2023) The impact of local financial development and corruption control on firm efficiency in Vietnam: evidence from a geoadaptive stochastic frontier analysis. *J Prod Anal* 60:203–226
- Haschka RE, Herwartz H, Struthmann P, Tran VT, Walle YM (2021) The joint effects of financial development and the business environment on firm growth: evidence from Vietnam. *J Comp Econ* 50:486–506
- Haschka RE, Wied D (2022) Estimating fixed-effects stochastic frontier panel models under “wrong” skewness with an application to health care efficiency in Germany. Available at SSRN https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4079660
- Ho H-A (2015) Business compliance with environmental regulations: evidence from Vietnam. *Vietnam Econ Annu Meet* 26:1–21
- Horrace WC, Parmeter CF, Wright IA (2024) On asymmetry and quantile estimation of the stochastic frontier model. *J Prod Anal* 61:19–36
- Huang T-H, Liu N-H, Kumbhakar SC (2018) Joint estimation of the Lerner index and cost efficiency using copula methods. *Empir Econ* 54:799–822
- Johnson NL, Kotz S, Balakrishnan N (1995) Continuous univariate distributions, vol 2, John Wiley & Sons
- Jondrow J, Lovell CK, Materov IS, Schmidt P (1982) On the estimation of technical inefficiency in the stochastic frontier production function model. *J Econom* 19:233–238
- Karakaplan MU, Kutlu L (2015) Handling endogeneity in stochastic frontier analysis. Available at SSRN 2607276
- Kumbhakar SC, Lovell CK (2003) Stochastic frontier analysis, Cambridge University Press
- Kumbhakar SC, Parmeter CF, Zelenyuk V (2020) Stochastic frontier analysis: foundations and advances. In Ray RG, Subhash C Chambers, Kumbhakar SC (eds) *Handbook of production economics*, Springer
- Kumbhakar SC, Schmidt P (2016) Editors’ introduction to the special volume “endogeneity problems in econometrics”. *J Econom* 190:209–211
- Kutlu L (2010) Battese-Coelli estimator with endogenous regressors. *Econ Lett* 109:79–81
- Lai H-P, Kumbhakar SC (2020) Estimation of a dynamic stochastic frontier model using likelihood-based approaches. *J Appl Econometr* 35:217–247
- Le V, Harvie C (2010) Firm performance in Vietnam: evidence from manufacturing small and medium enterprises. Department of Economics, University of Wollongong, Working Paper 04-10. <https://ro.uow.edu.au/commwkpapers/221>
- Le V, Vu X-BB, Nghiem S (2018) Technical efficiency of small and medium manufacturing firms in Vietnam: a stochastic meta-frontier analysis. *Econ Anal Policy* 59:84–91
- Li Q (1996) Estimating a stochastic production frontier when the adjusted error is symmetric. *Econ Lett* 52:221–228
- Møllgaard HP, Overgaard PB (2001) Market transparency and competition policy. *Rivista Polit Econ* 91:11–64
- Mutter RL, Greene WH, Spector W, Rosko MD, Mukamel DB (2013) Investigating the impact of endogeneity on inefficiency estimates in the application of stochastic frontier analysis to nursing homes. *J Prod Anal* 39:101–110
- Nelder JA, Mead R (1965) A simplex method for function minimization. *Comput J* 7:308–313
- Nghiem Tan L, Hau Long L, TRAN TVT (2021) Determinants of technical efficiency of microenterprises in Vietnam. *J Asian Financ Econ Bus* 8:829–838
- Nguyen DP, Ho VT, Vo XV (2018) Challenges for Vietnam in the globalization era. *Asian J Law Econ* 9:20180002
- Nguyen TT, Van Dijk MA (2012) Corruption, growth, and governance: private vs. state-owned firms in Vietnam. *J Bank Financ* 36:2935–2948
- Ortega MJR (2010) Competitive strategies and firm performance: technological capabilities’ moderating roles. *J Bus Res* 63:1273–1281
- O’Toole C, Newman C (2017) Investment financing and financial development: evidence from Vietnam. *Rev Financ* 21:1639–1674
- Papadopoulos A (2021) Measuring the effect of management on production: a two-tier stochastic frontier approach. *Empir Econ* 60:3011–3041
- Papadopoulos A (2022) Accounting for endogeneity in regression models using Copulas: a step-by-step guide for empirical studies. *J Econom Methods* 11:127–154
- Papadopoulos A, Parmeter CF (2023) The wrong skewness problem in stochastic frontier analysis: a review. *J Prod Anal* 61:1–14
- Papies D, Ebbes P, Van Heerde HJ (2017) Addressing endogeneity in marketing models. In Leeflang P, Wieringa J, Bijmolt T, Pauwels K (eds) *Advanced methods for modeling markets*, vol 1 of International series in quantitative marketing, 18, Basel: Springer, p 581–627
- Park S, Gupta S (2012) Handling endogenous regressors by joint estimation using Copulas. *Mark Sci* 31:567–586
- Parmeter CF, Racine JS (2013) Smooth constrained frontier analysis. In Chen X, Swanson N (eds) *Recent advances and future directions in causality, prediction, and specification analysis: essays in honor of Halbert L. White Jr*, New York: Springer, p 463–489
- Prokhorov A, Schmidt P (2009) Likelihood-based estimation in a panel setting: Robustness, redundancy and validity of copulas. *J Econom* 153:93–104
- Prokhorov A, Tran KC, Tsionas MG (2021) Estimation of semi- and nonparametric stochastic frontier models with endogenous regressors. *Empir Econ* 60:3043–3068
- Rand J, Tarp F (2012) Firm-level corruption in Vietnam. *Econ Dev Cult Change* 60:571–595
- Redmond W (2013) Three modes of competition in the marketplace. *Am J Econ Sociol* 72:423–446
- Reeb D, Sakakibara M, Mahmood IP (2012) From the editors: endogeneity in international business research. *J Int Bus Stud* 43:211–218
- Rioja F, Valev N (2004) Finance and the sources of growth at various stages of economic development. *Econ Inquiry* 42:127–140
- Sahut J-M, Teulon F (2022) The challenges of the transition to market economies for post-communist East Asian countries. *Post-Communist Econ* 34:283–292
- Shee A, Stefanou SE (2015) Endogeneity corrected stochastic production frontier and technical efficiency. *Am J Agric Econ* 97:939–952
- Siebert RB (2017) A structural model on the impact of prediscoversy licensing and research joint ventures on innovation and product market efficiency. *Int J Ind Organ* 54:89–124
- Simar L, Wilson PW (2009) Estimation and inference in cross-sectional, stochastic frontier models. *Econom Rev* 29:62–98
- Smith MD (2008) Stochastic frontier models with dependent error components. *Econom J* 11:172–192
- Son TVH, Coelli T, Fleming E (1993) Analysis of the technical efficiency of state rubber farms in Vietnam. *Agric Econ* 9:183–201

- Torii A (1992) Technical efficiency in Japanese industries. In Caves RE (ed) *Industrial efficiency in six nations*, Cambridge: MIT Press, p 31–119
- Tran KC, Tsionas EG (2013) GMM estimation of stochastic frontier models with endogenous regressors. *Econ Lett* 118:233–236
- Tran KC, Tsionas EG (2015) Endogeneity in stochastic frontier models: copula approach without external instruments. *Econ Lett* 133:85–88
- Tran TB, Grafton RQ, Kompas T (2008) Firm efficiency in a transitional economy: evidence from Vietnam. *Asian Econ J* 22:47–66
- Tsionas EG (2007) Efficiency measurement with the Weibull stochastic frontier. *Oxford Bull Econ Stat* 69:693–706
- Tsionas MG (2017) “When, where, and how” of efficiency estimation: Improved procedures for stochastic frontier modeling. *J Am Stat Assoc* 112:948–965
- Vu QN (2003) Technical efficiency of industrial state-owned enterprises in Vietnam. *Asian Econ J* 17:87–101
- Waldman DM (1982) A stationary point for the stochastic frontier likelihood. *J Econom* 18:275–279
- Wiboonpongse A, Liu J, Sriboonchitta S, Denoeux T (2015) Modeling dependence between error components of the stochastic frontier model using copula: application to intercrop coffee production in Northern Thailand. *Int J Approx Reason* 65:34–44